MICROBIAL REGULATION OF TRANSLATION

EDITED BY: Assaf Katz, Michael Ibba and Sebastian Andreas Leidel
PUBLISHED IN: Frontiers in Genetics

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88966-360-6
DOI 10.3389/978-2-88966-360-6

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

MICROBIAL REGULATION OF TRANSLATION

Topic Editors:
Assaf Katz, University of Chile, Chile
Michael Ibba, The Ohio State University, United States
Sebastian Andreas Leidel, University of Bern, Switzerland

Citation: Katz, A., Ibba, M., Leidel, S. A., eds. (2021). Microbial Regulation of Translation. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88966-360-6

Table of Contents

frontiers in Genetics

EDITORIAL
published: 23 November 2020
doi: 10.3389/fgene.2020.616946

Editorial: Microbial Regulation of Translation

Assaf Katz[1]*, *Sebastian A. Leidel*[2]* and *Michael Ibba*[3]*

[1] Programa de Biología Celular y Molecular, Instituto de Ciencias Biomèdicas, Facultad de Medicina, Universidad de Chile, Santiago, Chile, [2] Research Group for RNA Biochemistry, Department of Chemistry and Biochemistry, University of Bern, Bern, Switzerland, [3] Schmid College of Science and Technology, Chapman University, Orange, CA, United States

Keywords: translation, ribosome, tRNA, regulation, microorganisms

Editorial on the Research Topic

Microbial Regulation of Translation

Since the description of the operon model by Jacob and Monod during the late 1950s and early 1960s (Ullmann, 2010), the concept that the reading of genetic information must be a regulated process has been central to our understanding of biology. This is particularly true for microbes, which can adapt to an incredible variety of environments. Based on the research performed since the description of the operon, we have gained a deep understanding of the diverse strategies used by microbes to modulate the transcription of genetic information from DNA to RNA. In contrast, the mechanisms that regulate the translation of messenger RNAs into proteins has received less attention. The technical developments of the last decade now allow us to obtain detailed information on RNA folding (Rouskin et al., 2014; Aw et al., 2016) and modification (Linder et al., 2015; Lorenz et al., 2020) and the speed of translation (Subramaniam et al., 2013; Ingolia, 2014; Dai et al., 2016). This, in turn, allows us to scrutinize the functionality of translation components *in vivo*, providing unprecedented opportunities to study translation regulation. In this special issue of Frontiers in Genetics, "Microbial Regulation of Translation," we have assembled a series of articles that use diverse experimental approaches to study the regulation of translation in microbes.

Some of the papers in this issue focus on alterations of translation derived from changes in ribosome function and abundance. For instance, Pletnev et al. studied the physiological and molecular effects of mutating all genes known to methylate nucleotides of rRNA in *Escherichia coli*. While the mutation of some genes strongly impacts bacterial replication, others only lead to minor effects. Interestingly, with the exception of the *rsfM* mutation, most mutants exhibit defects in translation when the system is challenged by overexpression of exogenous genes, although some of these strains show only small effects under "normal" conditions. The article by Yoshida et al. is also related to changes in ribosome availability in *E. coli*. Nevertheless, this work focuses on the natural regulation of ribosome availability through hibernation and how this is coordinated with the abundance of RNA polymerase and its diverse sigma factors.

Other articles in this issue study the regulation of the initiation and elongation steps of translation. For instance, one article (Radío et al.) shows how ribosome profiling can be used to study the regulation of translation initiation by uORFs in *Trypanosoma cruzi*, a mechanism that accounts for regulation of almost 10% of the genes. Other articles instead discuss regulation of translation elongation. Leiva et al. shows how inactivation of $tRNA^{Gly}$ under oxidative stress may regulate translation elongation in *E. coli*, thereby changing protein synthesis only under specific environmental conditions. Using a different bacterial model, *Mycobacterium smegmatis*, Barth and Woychik analyze the role of the toxin-antitoxin system MazEF-ms in the regulation of translation. They found that MazEF-ms cleaves $tRNA^{Lys}$, thereby decreasing the speed of translation elongation at AAA Lys codons. Thus, when this toxin is activated, expression of AAA rich genes is decreased while expression of AAA poor genes is increased.

OPEN ACCESS

Edited by:
William C. Cho, QEH, Hong Kong

Reviewed by:
Malgorzata Kloc, Houston Methodist Research Institute, United States
Hyouta Himeno, Hirosaki University, Japan

***Correspondence:**
Assaf Katz
askatz@uchile.cl
Sebastian A. Leidel
sebastian.leidel@dcb.unibe.ch
Michael Ibba
ibba@chapman.edu

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *13 October 2020*
Accepted: *26 October 2020*
Published: *23 November 2020*

Citation:
Katz A, Leidel SA and Ibba M (2020) Editorial: Microbial Regulation of Translation. Front. Genet. 11:616946. doi: 10.3389/fgene.2020.616946

In addition to the paper by Barth and Woychik, two reviews in this issue analyze the effects of diverse toxins on the translation machinery. Jurenas and van Melderen focus on the ability of many type II toxin-antitoxin systems to interfere with translation. These systems are composed of a toxin that may inhibit central processes of a cell and its antitoxin, a labile protein that inhibits the activity of the toxin. The authors discuss how most of these toxins' targets are components of the translation apparatus, including mRNA, tRNA, ribosomes, and translation factors. They further propose that the huge variability of these systems derives from low selective pressure on bacteria to maintain them, and the high selective pressure on the toxin-antitoxin systems to diverge from similar systems allowing lateral transfer to organisms carrying similar toxin-antitoxins. Thus, they propose these are "selfish genes" that usually give little advantage to bacteria. The second review about toxins uses a very different approach. The text written by Travin et al. is focused on ribosomally synthesized and post-translationally modified peptides (RiPPs). Similar to type II toxin-antitoxin systems, many of the RiPPs that have been described target diverse components of the translation apparatus. Nevertheless, in contrast to type II toxin-antitoxin systems, these compounds are not targeted to inhibit the self-translation machinery, but that of competitor organisms. Travin et al. provide an in-depth description of several of these compounds from a structural and functional perspective and highlight strategies to screen for pathways that produce new varieties of these compounds and their potential pharmacological use.

Finally, two papers in this issue use an unbiased screening to identify patterns in regulation of gene expression. Gummesson et al. describe the use of spike-in normalized RNA sequencing, applying it to *E. coli* cells subjected to valine-induced isoleucine starvation. In addition to providing valuable biological data, they show how changes in the total RNA levels may interfere with typical normalization protocols used to analyze transcriptomic data. Finally, Zhao et al. describe the usage of combined proteomics and genomics to study the relation between the organization of genes in operons on their cellular concentrations. They find interesting correlations, particularly for operons coding genes from a single complex or metabolic pathway.

In total, the collection of papers included in this issue represents the enormous variety of approaches and findings of an area that has been invigorated by the developments of the last decade.

AUTHOR CONTRIBUTIONS

AK wrote the paper with contributions from SL and MI. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by Fondo Nacional de Desarrollo Científico y Tecnológico [1191074 to AK]; National Institutes of Health Grant [GM65183 to MI]; and the Swiss National Science Foundation [NCCR RNA & Disease to SL].

REFERENCES

Aw, J. G., Shen, Y., Wilm, A., Sun, M., Lim, X. N., Boon, K. L., et al. (2016). *In vivo* mapping of eukaryotic RNA interactomes reveals principles of higher-order organization and regulation. *Molecular Cell* 62, 603–617. doi: 10.1016/j.molcel.2016.04.028

Dai, X., Zhu, M., Warren, M., Balakrishnan, R., Patsalo, V., Okano, H., et al. (2016). Reduction of translating ribosomes enables *Escherichia coli* to maintain elongation rates during slow growth. *Nat. Microbiol.* 2:16231. doi: 10.1038/nmicrobiol.2016.231

Ingolia, N. T. (2014). Ribosome profiling: new views of translation, from single codons to genome scale. *Nat. Rev. Genet.* 15, 205–213. doi: 10.1038/nrg3645

Linder, B., Grozhik, A. V., Olarerin-George, A. O., Meydan, C., Mason, C. E., and Jaffrey, S. R. (2015). Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat. Methods* 12, 767–772. doi: 10.1038/nmeth.3453

Lorenz, D. A., Sathe, S., Einstein, J. M., and Yeo, G. W. (2020). Direct RNA sequencing enables m6A detection in endogenous transcript isoforms at base-specific resolution. *RNA* 26, 19–28. doi: 10.1261/rna.072785.119

Rouskin, S., Zubradt, M., Washietl, S., Kellis, M., and Weissman, J. S. (2014). Genome-wide probing of rna structure reveals active unfolding of mRNA structures *in vivo*. *Nature* 505, 701–705. doi: 10.1038/nature12894

Subramaniam, A. R., Pan, T., and Cluzel, P. (2013). Environmental perturbations lift the degeneracy of the genetic code to regulate protein levels in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 110, 2419–2424. doi: 10.1073/pnas.1211077110

Ullmann, A. (2010). *Escherichia coli* and the French school of molecular biology. *EcoSal Plus* 4, 1–13. doi: 10.1128/ecosalplus.1.1.1

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

ORIGINAL RESEARCH
published: 24 May 2019
doi: 10.3389/fgene.2019.00473

Multifaceted Stoichiometry Control of Bacterial Operons Revealed by Deep Proteome Quantification

Jing Zhao[1], Hong Zhang[1], Bo Qin[2], Rainer Nikolay[2], Qing-Yu He[1], Christian M. T. Spahn[2] and Gong Zhang[1]**

[1] *Key Laboratory of Functional Protein Research of Guangdong Higher Education Institutes, Institute of Life and Health Engineering, Jinan University, Guangzhou, China,* [2] *Institut für Medizinische Physik und Biophysik, Charité – Universitätsmedizin Berlin, Berlin, Germany*

More than half of the protein-coding genes in bacteria are organized in polycistronic operons composed of two or more genes. It remains under debate whether the operon organization maintains the stoichiometric expression of the genes within an operon. In this study, we performed a label-free data-independent acquisition hyper reaction monitoring mass-spectrometry (HRM-MS) experiment to quantify the *Escherichia coli* proteome in exponential phase and quantified 93.6% of the cytosolic proteins, covering 67.9% and 56.0% of the translating polycistronic operons in BW25113 and MG1655 strains, respectively. We found that the translational regulation contributes largely to the proteome complexity: the shorter operons tend to be more tightly controlled for stoichiometry than longer operons; the operons which mainly code for complexes is more tightly controlled for stoichiometry than the operons which mainly code for metabolic pathways. The gene interval (distance between adjacent genes in one operon) may serve as a regulatory factor for stoichiometry. The catalytic efficiency might be a driving force for differential expression of enzymes encoded in one operon. These results illustrated the multifaceted nature of the operon regulation: the operon unified transcriptional level and gene-specific translational level. This multi-level regulation benefits the host by optimizing the efficiency of the productivity of metabolic pathways and maintenance of different types of protein complexes.

Keywords: operons, proteome quantification, HRM-MS, multifaceted stoichiometry control, mass-spectrometry, DIA, translation

OPEN ACCESS

Edited by:
Michael Ibba,
The Ohio State University,
United States

Reviewed by:
Jesse Rinehart,
Yale University, United States
Michelle Gibbs,
The Ohio State University,
United States

****Correspondence:***
Christian M. T. Spahn
christian.spahn@charite.de
Gong Zhang
zhanggong-uni@qq.com

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *24 February 2019*
Accepted: *01 May 2019*
Published: *24 May 2019*

Citation:
Zhao J, Zhang H, Qin B, Nikolay R, He Q-Y, Spahn CMT and Zhang G (2019) Multifaceted Stoichiometry Control of Bacterial Operons Revealed by Deep Proteome Quantification. Front. Genet. 10:473. doi: 10.3389/fgene.2019.00473

INTRODUCTION

An operon is a cluster of genes transcribed in a single mRNA. This principle is conserved across bacterial and archaeal genomes, as well as mitochondria and chloroplast (Wolf et al., 2001; Price et al., 2005; Zheng et al., 2005). Operons are also found in virus and some lower eukaryotes, including yeasts, nematodes, and insects (Blumenthal, 2004; Ben-Shahar et al., 2007; Osbourn and Field, 2009; Pi et al., 2009; Gordon et al., 2015). In a typical bacterial genome, more than half of the protein-coding genes are organized in multigene operons. A classical bacterial operon generates an mRNA strand with polycistronic structure containing multiple coding sequences and are translated together in the cytoplasm. These genes are often of related functions, for example, to build a protein

complex or to participate in one metabolic pathway. Therefore, grouping related genes as operons under the control of a single promotor is often thought to simplify the regulation of gene expression for rapid adaptation to environmental changes.

An intuitive presumption of the operon organization is to maintain stoichiometry of the gene products. It was argued that co-regulation could be evolved by merging two independent genes in proximity together under the control of the same promoter, to reduce the control complexity (Lawrence and Roth, 1996; Osbourn and Field, 2009), Li et al. (2014) measured protein synthesis rates by using ribosome profiling and implied that the synthesis rates quantitatively might reflect the stoichiometry of the protein complexes. Studies showed that an operon with one complex promoter might be better than two independent promoters; organization of genes in operons substantially reduces the shortfall in production of complex-forming individual proteins (Iber, 2006; Osbourn and Field, 2009). However, recent advances of omics techniques raised counter-arguments. A transcriptome-level study revealed that certain adjacent genes within one operon are not similarly transcribed in *M. pneumoniae*. In half of the polycistronic operons, genes exhibited a decaying expression according to its rank in the operon, which is termed "staircase-like decay behavior" (Guell et al., 2009). Considering the widespread post-transcriptional regulations including translational control and protein turnover (Schwanhausser et al., 2011), it is still under intensive debate whether this "staircase-behavior" influences the protein abundance (Maier et al., 2011; Schmidt et al., 2011; Arike et al., 2012).

Theoretically, proteins in a complex should follow the stoichiometry, while the proteins involved in the same pathway may need differential expression controls (Guell et al., 2009). For example, the enzymes in various amino acids synthesis pathways are regulated in single-input modules (SIMs). A series of such enzymes are successively expressed in one operon (Zaslaver et al., 2004; Seshasayee et al., 2006). Meanwhile, the different catalytic kinetics of these enzymes determines that these enzymes should not be expressed at the same quantity (Zaslaver et al., 2004). These genes tend to duplicate to evolve a larger gene regulatory network (Teichmann and Babu, 2004), indicating their regulation is less stringent, and an operon arrangement might be unnecessary. Therefore, a more detailed proteome-wide and quantitative investigation is necessary to discover the scope and impact of the operons in gene expression regulation.

A method capable to assess a quantification of the proteome should be used in this case. Although stable isotope labeling methods are more accurate than label-free mass-spectrometry (MS) methods (Arike et al., 2012), the isotopes may affect the physiology of the bacteria (Xie and Zubarev, 2015). The isotope labeling is more suitable for comparative quantification of multiple samples than estimating abundance of the proteins within one sample (Neilson et al., 2011). Therefore, label-free MS methods should be used. Arike et al. (2012) compared three label-free quantification methods (iBAQ, emPAI, and APEX) and found a staircase-like protein expression in most of the transcription units, and found high correlation abundances between some well-known complex subunits. In contrast, Schmidt et al. (2011) found only 5% "staircase behavior" for *L. interrogans* operons on the proteome level. These contradictory results reflected the cons of these label-free MS approaches: the technical variations and relatively low number of quantified proteins restricted the accurate and in-depth coverage of operon-controlled genes.

In this work, we set out to tackle these problems by employing a highly accurate label-free method, DIA (data-independent acquisition) (Purvine et al., 2003), to obtain quantification of the proteins constituting the *Escherichia coli* proteome with a high coverage and high accuracy. DIA is a MS-based proteomics method used in peptide quantification, in which all ions within a selected m/z range are fragmented and analyzed in a second stage of tandem mass spectrometry (Law and Lim, 2013). Although not suitable for discovery-based applications, DIA provides accurate peptide quantification without being limited to profiling predefined peptides of interest (Chapman et al., 2014; Doerr, 2015). This allowed us to investigate the protein abundances within operons and thus to interrogate the possible stoichiometry in operons of different functions.

MATERIALS AND METHODS

MS Sample Preparation

Escherichia coli K-12 sub-strains BW25113 and MG1655 were cultivated on glucose M9 minimal medium at 37°C in flasks to mid-exponential phase (OD_{600} = 0.6) and then harvested in 45 mL volume, immediately cooled in ice water, and then centrifuged at 10,000 × *g* for 5 min. The pellet was washed once with PBS, centrifuged at 10,000 × *g* for 5 min again. Pellet was re-suspended on ice with lysis buffer (5 M urea/2 M thiourea in 10 mM HEPES, pH 8.0), and were sonicated and centrifuged at 17,000 × *g* for 30 min in a table-top centrifuge to remove cell debris. Supernatant was collected, and protein concentrations were determined with a Bradford Protein Assay (Bio-Rad Protein Assay Dye Reagent Concentrate, Cat. #500-0006).

For proteome analysis, we employed in-solution protein digestion with a filter-aided sample preparation (FASP) method (Zhang et al., 2017). 1 mg of protein was subjected to reduction (8 M urea and 50 mM DTT at 37°C, 1 h), followed by alkylation with 100 mM iodoacetamide (IAA) in dark at room temperature for 30 min. The solution was transferred to the 30 kDa ultracentrifuge filters (Millipore). Proteins were washed with 8 M urea, and four sequential buffer changes were performed using 50 mM TEAB, respectively. Trypsin (Promega) was then added into the filter at a mass ratio of 1:20 for Proteins digested in 130 μL 50 mM TEAB at 37°C for 12 h. The released peptides were collected by centrifugation and dried with a cold-trap speed vacuum.

MS Experiments

One microgram of sample abovementioned peptides was analyzed on a C18 column (50 μm × 15 cm, 2 μm, Thermo Fisher) by using an EASY-nLC 1200 UHPLC connected to an

Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific). The peptides with the iRT-standard (1/10 by volume, Biognosys, HRM Calibration Kit: Ki-3003) were separated by a linear gradient from 6 to 30% ACN with 0.1% formic acid at 270 nL/min for 100–130 min and linearly increased to 90% ACN in 20 min. For the data-dependent acquisition (DDA), the source was operated at 2.0 kV. The DDA scheme included a full MS survey scan from *m/z* 400 to *m/z* 1500 at a resolution of 60,000 FWHM with AGC set to 4E5 (maximum injection time of 50 ms), followed by MS/MS scans at a resolution of 15,000 FWHM with AGC set to 5E4 (maximum injection time of 30 ms), data-dependent mode was set to top speed. Isolation window was 1.6. Dynamic exclusion was set to 90 s with a 10 ppm tolerance around the selected precursor. For the DIA hyper reaction monitoring (HRM-MS), individual tryptic peptide samples were mixed with the iRT-standard (1/10 by volume) and analyzed by the same method as DDA used. The method consisted of a full MS1 scan at a resolution of 60,000 FWHM from *m/z* 350 to *m/z* 1,200 with AGC set to 4E5 (maximum injection time of 30 ms) followed by 40 non-overlapping DIA windows acquired at a resolution of 30,000 FWHM with AGC set to 5E5 (maximum injection time of 50 ms), cycle time, 3.28 s. The MS/MS isolation windows were listed in **Supplementary Table S1**. For comparison, standard DDA MS experiment was performed as above. All MS raw data have been deposited in iProX with accession number: IPX0001095000 and ProteomeXchange with identifier PXD010126.

Spectral Library Generation

To generate the spectral library, three DDA measurements of the mixed samples were performed. Raw DDA datasets were searched against a combined database of the NCBI database of *Escherichia coli* str. K-12 (GCF_000005845.2_ASM584v2, 4140 entries) and the iRT standard peptides sequence using the Sequest HT (Proteome Discoverer v2.1) local server. Common contaminants in the database included trypsin and keratins. Precursor and product ion spectra were searched at an initial mass tolerance of 10 ppm and fragment mass tolerance 0.02 Da, respectively. Tryptic cleavage was selected, and up to two missed cleavages were allowed. Carbamidomethylation on cysteine (+57.021 Da) was set as a fixed modification, and oxidation (+15.995 Da) on methionine was assigned as a variable modification. A target-decoy-based strategy was applied to control peptide and protein false discovery rates (FDRs) at lower than 1%. Confident protein identifications should suit the following criteria: (1) protein level FDR $\leq$ 1%; (2) unique peptides $\geq$ 1 or 2; (3) peptide length $\geq$ 6 or 7 aa. The search result was exported in a pdResult file format containing the annotation of precursors and fragment ions and their exact retention times. The pdResult file was then imported into Spectronaut Pulsar 11 (Biognosys) to generate the spectral library used for HRM-MS data analysis, which yielded 14608 unique peptide sequences in 2041 protein groups with BW25113, and 8822 unique peptide sequences in 1607 protein groups with MG1655. A subset of identified peptides was used in library creation as modification parameter was set none. The generated spectral libraries were exported from Spectronaut as in **Supplementary Table S2**.

Protein Identification and Quantification

The DIA data were then analyzed with Spectronaut Pulsar 11 with the spectral library, which is a mass spectrometer vendor independent software for SWATH/DIA data analysis. Raw data were analyzed according to the user manual of the software. Default settings were setup for protein identification and peak area calculation. Raw data were converted into HTRMS files and imported to Spectronaut Pulsar 11 by choosing the matched database fasta file and spectral library, with the default settings of the Spectronaut Pulsar 11: (1) Calibration: calibration mode, automatic; iRT calibration strategy, non-liner iRT calibration. (2) Identification: decoy limit strategy, dynamic; decoy method, mutated; machine learning, per run; precursor q-value (peptide FDR) cutoff, 0.01; protein q-value (protein FDR) cutoff, 0.01; *p*-value estimator, kernel density estimator. (3) Workflow: default labeling type, label; profiling strategy, none; unify peptide peaks, false. (4) Quantification: interference correction, true; major(protein) grouping, by protein-group id; major group quantity, mean peptide quantity; minor (peptide) grouping, by stripped sequence; minor group quantity, mean precursor quantity; minor group top n, true; min, 1; max, 3; quantity MS-level, MS2; quantity type, area; data filtering, *q*-value; cross run normalization, true; row selection, *q*-value sparse; normalization strategy, local normalization. (5) Reporting: scoring histograms, true; pipeline report schema, protein quant; pipeline reporting unit, experiment. (6) Protein inference: protein inference workflow, automatic. (7) Data extraction: MS1 mass tolerance strategy, dynamic; correction factor, 1; MS2 mass tolerance strategy, dynamic; correction faction, 1. (8) Post analysis: differential abundance grouping, major group (quantification settings); smallest quantitative unit, precursor ion (summed fragment ions); use top n selection, false. (9) Retention time were used to assist identification. XIC extraction: XIC RT extraction window, dynamic; correction factor, 1. After peak extraction and area calculation were performed, the result was exported as the table format for further quantification analysis in Microsoft Excel. All MS raw data, Proteome Discoverer report (*.msf file) and the constructed spectra library have been deposited in iProX with accession number: IPX0001095000 and ProteinXchange with identifier PXD010126. The relationship of submitted raw data are shown in **Supplementary Data**.

Protein abundances was calculated by using Spectronaut Pulsar protein pivot report, those proteins quantified by Spectronaut pulsar but not identified by Proteome Discoverer and not met the confident protein identifications criteria were removed. The abundances of identified proteins were calculated as follow procedure. Supposed that top 500 abundant proteins of *E. coli* can represent total protein copy numbers of a cell. Concentration of HRM-MS protein copies per cell was calculated based on the means of 500 most abundant protein quantities computed by other three label-free methods (APEX, iBAQ, PAI) downloaded from Arike et al. (2012). HRM-MS intensity could

be converted to protein copies per cells by coefficient k, which is defined by the following formula.

B_i is copy numbers of the gene in iBAQ dataset.

P_i is copy numbers of the gene in emPAI dataset.

A_i is copy numbers of the gene in APEX dataset.

D_i is correspondence protein intensities of the gene quantified in HRM-MS dataset.

$$k = \frac{\sqrt[n]{\prod_{i=0}^{n} \frac{(Bi+Pi+Ai)}{3}}}{\sqrt[n]{\prod_{i=0}^{n} Di}} \quad (1)$$

The amount of individual proteins was calculated as the product of conversion coefficient k to their intensity in the HRM-MS measured sample.

$$Protein\ copy\ number = k \times Di \quad (2)$$

The calculations were performed by in-house generated python scripts. All scripts used in this study can be downloaded in the **Supplementary Materials** (**Supplementary Scripts**).

Coefficient of Variation of Protein Abundance in the Operons

Coefficient of variation (CV), which is defined as the ratio of the standard deviation to the arithmetic mean. Standard deviation is normalized by $n - 1$ by default (n is sample size). The CV of proteins within one operon is defined as the ratio of the standard deviation of protein quantities within this operon to the arithmetic mean of all protein quantities within this operon. For multi-gene operons (protein numbers ≥ 2), CV was calculate as follows:

$$CV = \frac{\sqrt[2]{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}}{mean} \times 100\% \quad (3)$$

where x_i is the abundance of the i-th gene in this operon. To be noted, the CV calculation was only performed within one operon, not across the operons.

The CV of the protein half-life in the operons were calculated in the same way. Protein half-life time in the M9 minimal medium was from our previous work (Zhong et al., 2015).

Data Randomization

To compare with the real operon CV level if protein abundances in operons have stoichiometry control, we reshuffled the protein quantities detected in the polycistronic operons (846) randomly to each protein ID, the generated dataset was used as randomized negative control. Randomized protein quantities in "2-/3-/4-/≥5-protein" operons were extracted from this randomized negative control data.

Operon and Gene Ontology (GO) Analysis

Operon library of *Escherichia coli* str. K-12 were downloaded from the DOOR[2] database (Mao et al., 2014) (NC_000913). Protein GI numbers were converted to proper identifiers by DAVID Gene Accession Conversion Tool (Huang et al., 2007). The quantified proteins were integrated to the operon data. The PANTHER Version 13.0 (released 2017-11-12[1]) (Mi et al., 2017) was used to perform the GO overrepresentation analysis with the significance threshold of 0.01, the quantified proteins of *Escherichia coli* in our work was selected as the background proteome, the Fisher's Exact test was used to obtain p-values and 'GO slim' category were used. The protein subcellular localization data of *E. coli* was downloaded from EcoProDB (Yun et al., 2007).

Complex and Pathway Classification

The operon contains more than or equal to two genes were called polycistronic operons. Among polycistronic operons, those ≥90% genes in operon encoded subunits of one protein complex is selected and classified as "Complex" group, others were classified as "Pathway" group.

Physical and Chemical Features of Proteins

The protein lengths in amino acids were obtained from the NCBI of *Escherichia coli* str. K-12 (GCF_000005845.2_ASM584v2, 4140 entries). Information of the hydrophobicity was calculated by Gravy Calculator[2]. In addition, the isoelectric point, protein length, instability and hydrophobicity distribution were calculated by using python 2.7 scripts and Biopython libraries.

Experimental Design and Statistical Rationale

To increases the precision of protein expression measurements of the entire *E. coli* proteome quantification, two biological replicates of BW25113 and MG1655 each were cultured in M9 minimal medium to mid-exponential phase and were harvest, then processed to HRM-MS analysis independently. iRT-standard (Biognosys, HRM Calibration Kit) was added to the peptides with 1/10 by volume. The peptide mixture of two biological replicates of each strain were used and performed LC-MS for three times for spectral library creation. Proteome Discoverer 2.1 and Spectronaut Pulsar 11 were used to generated spectral library, and Spectronaut Pulsar 11 was used to quantify the protein groups with little modified parameters. Kolmogorov–Smirnov test (KS-test) were used to compare the distributions between CVs at transcriptome, translatome and protein level, and Mann-Whitney U-test were used to compare the difference between complex and pathway operons.

mRNA Sequencing

E. coli strain BW25113 was cultivated on glucose M9 minimal medium at 37°C in flasks to mid-exponential phase (OD_{600} = 0.6) with 100 μg/mL chloramphenicol added 15 min before harvest, then the cells were centrifuged at 5,000 × g for 10 min at 4°C, followed by thrice washed with pre-chilled PBS. Cell pellet was then re-suspended in 6 mL pre-chilled sucrose-buffer solution [16 mM Tris (pH 8.1) supplemented with 0.5 M RNase-free sucrose, 50 mM KCl, 8.75 mM

[1]http://pantherdb.org/

[2]http://www.gravy-calculator.de

EDTA, 100 μg/mL chloramphenicol, 12.5 mg/mL lysozyme] and gently stirred for 5 min on ice. Then the cells were centrifuged 5,000 × *g*, 10 min. Total RNA was extracted by Trizol method, and mRNA-seq libraries were prepared using standard MGIEasy^{TM} mRNA Library Prep Kit V2 following the manufacturer's protocol. Sequencing was performed on a BGISEQ-500 sequencer for 50 cycles, single-ended mode. This dataset was deposited in the GEO database under the accession number GSM3489376, GSM3489377.

Analysis of Sequencing Data

The RNA-seq dataset of *E. coli* strain MG1655 were obtained from Haft et al. (2014) (GEO accession number: GSM1360030, GSM1360031, GSM1360042, GSM1360043) and Bartholomaus et al. (2016) (SRA accession number: SRR2016457). The datasets of strain BW25113 were generated as described above. For all datasets, adapters was trimmed from the reads. Reads were mapped to coding sequence of *E. coli* reference genome (GenBank: U00096) using FANSe3 algorithm (Liu et al., 2017) with the parameters -E3 -S10 –indel. Genes with at least 10 mapped reads were considered quantifiable (Bloom et al., 2009). The expression levels were estimated in rpkM.

RESULTS

Near-Complete *E. coli* Cytosolic Proteome Quantification Using HRM-MS

To assess the quantification power and reproducibility of the HRM-MS method, we performed two biological replicates of *E. coli* total soluble proteins of strain BW25113 and MG1655. When using the previous identification criteria (at least one unique peptide, peptide length ≥6 amino acids) (Arike et al., 2012), our HRM-MS results quantified 1951 and 1571 proteins in these two strains, respectively. The two biological replications identified almost identical proteins, demonstrating high robustness and reproducibility (**Figure 1A**). Two replicates quantified almost the same proteins: only a few proteins were quantified only in one replicate (**Figure 1A**). Under the stringent criteria as two unique peptides and at least seven amino acids peptide length. Even under the stringent criteria, we still quantified 1675 and 1252 proteins with a high reproducibility (**Figures 1B,C** and **Supplementary Table S3**). The number of quantified soluble proteins was almost doubled when compared to the previous results quantified by other methods (1021 proteins for APEX, 1183 for IBAQ and 1138 for emPAI) (Arike et al., 2012). To rule out the difference of the instruments, we performed DDA MS experiments for the two strains in the same Orbitrap Fusion Lumos instrument. Proteins were identified under the stringent criteria and quantified using iBAQ method. The DDA MS quantified 1520 and 1482 proteins for BW25113 and MG1655 strains, respectively, comparable with the HRM-MS experiments. However, the correlation coefficients of the iBAQ quantification of two biological replicates were 0.932 and 0.939, respectively (**Figures 1D,E**), lower than the HRM-MS (R = 0.982 and 0.988 for the two strains, respectively). Each identified protein was covered by 19.20 and 15.45 peptides in average in two strains, covering 31.56% and 24.87% of the amino acid sequences, respectively (**Figure 1F**), which is higher than the typical peptide coverage of human proteome MS experiments (single search engine, up to ∼20% coverage) (Zhao et al., 2017).

We previously revealed 2922 genes which are being translated in the *E. coli* grown in the same condition using ribosome profiling (Bartholomaus et al., 2016). Using HRM-MS method with stringent criteria, we quantified 55.2% of these translating genes in this work. We next analyzed the possible chemical and physical features of the translated but unquantified proteins in this work. The unquantified proteins are significantly more alkalic, less stable, shorter and more hydrophobic (**Supplementary Figure S1**). These are general factors that decreases the visibility of these proteins in shotgun MS experiments. Since our experiments were not optimized for membrane proteins, which are more prone to aggregate during the protein extraction, these proteins are expected to be less detected in the MS. Notably, we quantified 93.6% translating cytosolic proteins, showing a near-complete quantification of the soluble proteins.

Considering the advantage of the HRM-MS, all the subsequent analysis was based on the HRM-MS under the stringent criteria.

Operons Tend to Unify Gene Expression in General

The high coverage of proteome quantification allowed us to make an in-depth investigation of protein abundances in operons. Indeed, our quantification covered 67.9% and 56.0% of the translating polycistronic operons in BW25113 and MG1655, respectively. We calculated the coefficient of variation (CV, %) of proteins abundances within each operon. To compare with the real operon CV level, we random redistribution the protein quantities quantified in the experiment as randomized negative control. Smaller CV represent the unified expression level of the proteins within one operon. The mean of real CV was significantly smaller than that of randomized negative control data (Mann–Whitney U-test, $p = 5.05 \times 10^{-10}$ and $p = 3.40 \times 10^{-5}$ for BW25113 and MG1655, respectively) (**Figure 2**). The median CV of quantified proteins were also smaller than the randomized control (Kruskal–Wallis H-test, $p = 1.44 \times 10^{-8}$ and $p = 1.18 \times 10^{-6}$ for BW25113 and MG1655, respectively). As a positive control, the trend to unified expression is also valid in transcription (an operon is transcribed as an entire mRNA) (**Supplementary Figure S2**). This indicated that most polycistronic operons were co-regulated, and the stoichiometry balance of protein abundances within polycistronic operons exists in general, although with exceptions. These results were in general consistent with previous studies (Ishihama et al., 2008; Arike et al., 2012).

Functional Enrichment of the Operon Stoichiometry Control

We then examined if the number of genes per operon could affect the stoichiometry of proteins encoded within operons in both BW25113 and MG1655 strains. We found all the medians of subgroups were lower than the randomized negative control

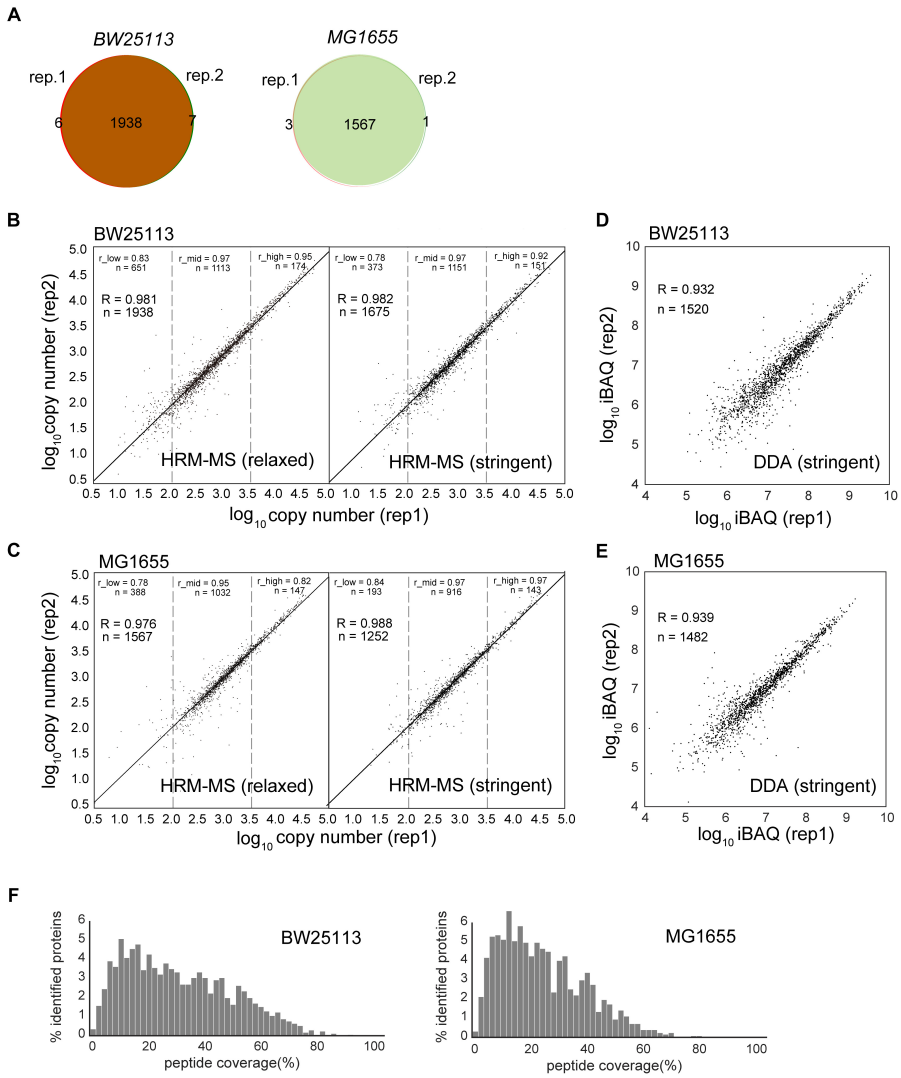

FIGURE 1 | Comparison of protein abundances obtained by different label-free quantification methods. **(A)** Quantified protein numbers in two biological replications of HRM-MS method of two strains. **(B,C)** The quantification reproducibility of HRM-MS method in relaxed and stringent criteria of two *E. coli* strains, respectively. R is the Pearson correlation coefficient. The protein abundance range was divided into low (<2.0 log scale), mid (2.0–3.5 log scale) and high (>3.5 log scale) sections. r_low, r_mid and r_high are the Pearson correlation of the proteins in these sections. **(D,E)** The reproducibility of DDA MS experiment of the two strains. Proteins were quantified using iBAQ method. **(F)** The peptide coverage of the HRM-MS-identified proteins in two *E. coli* strains, respectively.

(RD), which was in accordance with our abovementioned results (**Figure 3A**), indicating operon expression regulation exists in general. We continue to divide operons into two subgroups by their functions.

Next, each group was separated into two subgroups by the median CV of randomized data, the high CV subgroups (H2-H5, higher than median of randomized data) and the low CV subgroups (L2–L5, lower than median of randomized data). Gene ontology (GO) overrepresentation analysis was performed for each subgroup both in both strains (**Figures 3B,C**, see details in **Supplementary Table S5**). Most low CV subgroups (L2–L5) showed functional enrichment against the quantified proteins as background. The L2 group was overrepresented in almost all metabolic activities, and the L5 group was highly enriched in complexes and structural molecules in both *E. coli* strains. This provided a hint that the stringent stoichiometry control might be important for the efficient assembly of protein complexes. In contrast, there are little significant functional enrichment of GO terms enriched in H2–H5 subgroups in BW25113, and only the GO terms enriched in H2 and H5 subgroup in MG1655. These results indicated that the large CV of most of these operons might be caused by experimental error.

Since the operons in L2–L5 subgroups were enriched in metabolic pathways and complexes, we specifically divide these operons into "Complex" and "Pathway" groups. The group "Complex" are the operons encoding proteins for the same protein complex. The group "Pathway" are the operons encoding proteins involving in the same metabolic pathway. Similar to

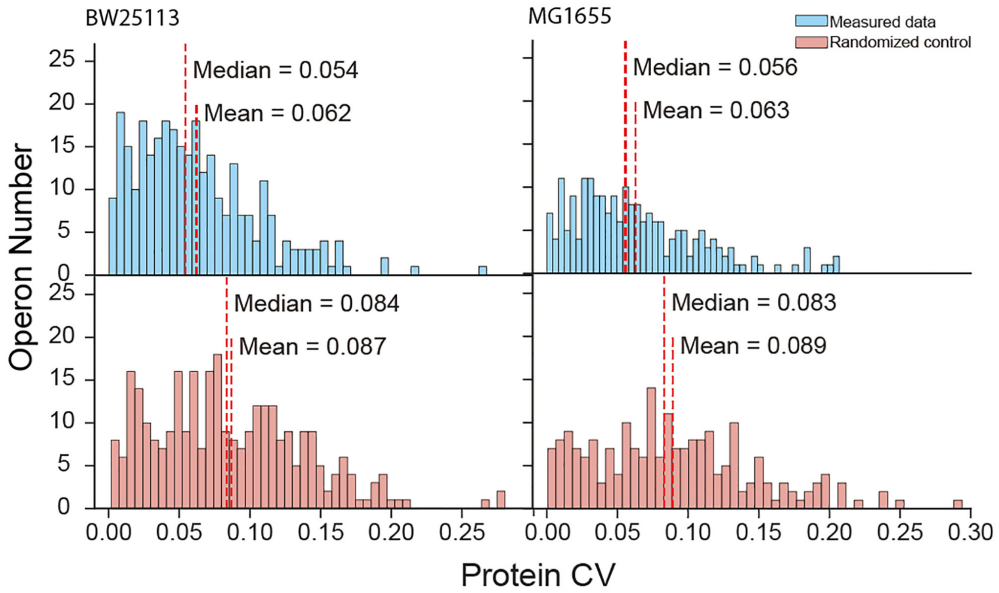

FIGURE 2 | Protein coefficient of variation (CV) within operons of measured data and randomized negative control.

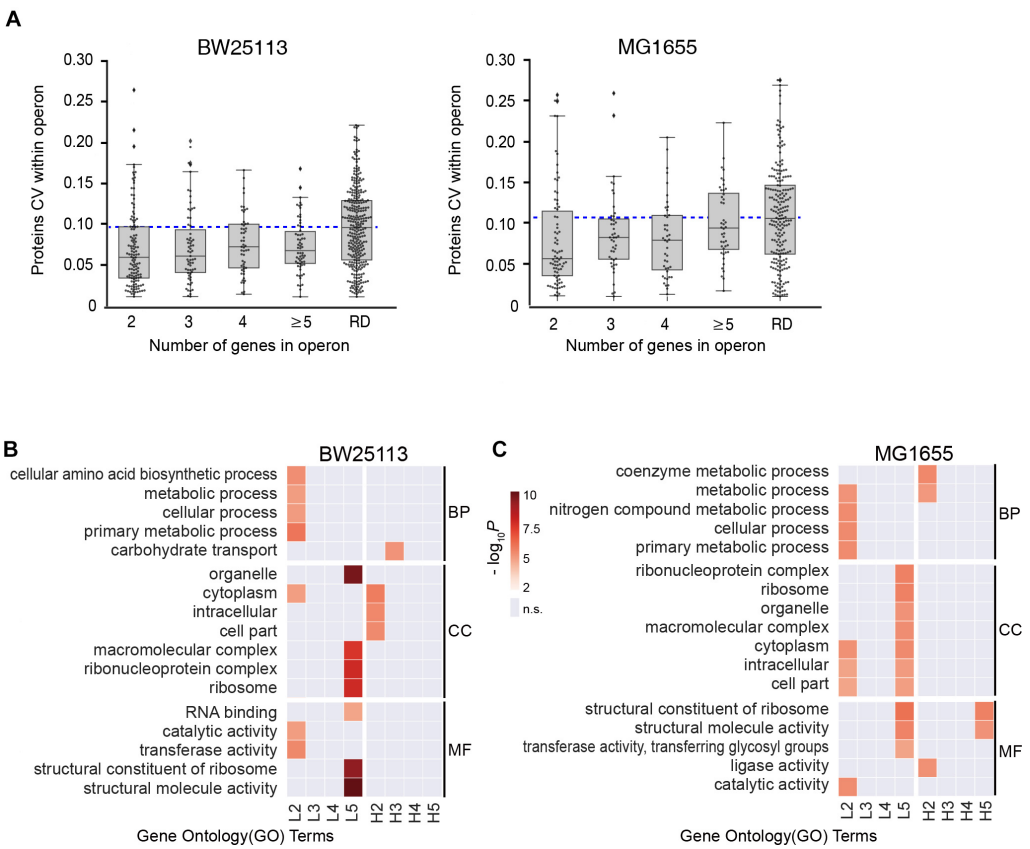

FIGURE 3 | Functional-dependence of operon stoichiometry control. **(A)** Protein CVs within operons, categorized according to the number of genes in operon. RD, randomized data. Blue dashed line represents the median of the randomized data. **(B,C)** Gene ontology (GO) overrepresentation of the operons with lower CV than the randomized median (L2–L5) and higher CV than the randomized median (H2–H5). The number refer to the number of genes in operon. GO terms with $P > 0.001$ were considered insignificant and marked as gray. BP, biological process; CC, cell component; MF, molecular function. **(B)** BW25113 and **(C)** MG1655 strain.

the **Figure 3A**, we performed randomization for each group multiple times for robustness. In almost all cases the 2-/3-/4-/≥5-protein "Complex" and "Pathway" operons exhibited lower CVs than the corresponding randomized data (**Supplementary Figure S3**), indicating that the stoichiometry is still maintained for the protein complexes and metabolic pathways in a certain extent, which is consistent with the traditional hypothesis.

However, there seems to be a trend that shorter operons (containing lower number of genes) possess lower CV within operon (**Figure 3A**), suggesting a length-dependent stoichiometry control. Since the mRNA of an entire operon is transcribed as a unit, the abovementioned phenomenon opened a question of the origin of the length-dependent stoichiometry control.

Length-Dependence of Stoichiometry Control Is Functional-Dependent

Linear regression analysis was performed to calculate the CVs at RNA and protein levels, within operon that encodes proteins forming complexes or involving in the metabolic pathways (**Figure 4A** and **Supplementary Figure S4**). Since the mRNA of one operon is transcribed as one unit, the CV within operon at RNA level is much lower than the randomized control, as expected (**Figure 4A** and **Supplementary Figure S4**). However, the "Pathway" operons did not show length dependence at protein level, while significant and positive correlation of protein CV within operon versus length were observed in "Complex" groups in both BW25113 and MG1655 (P-value of BW25113 = 0.0004, P-value of MG1655 = 0.022). In contrast, among the "Pathway" groups, those operons exhibited similar CV distribution regardless of their lengths (regression $P > 0.05$) (**Figure 4A**, "Protein" plots). This reflects necessity that the proteins operating in a pathway need to be more independently tuned and thus do not have to follow the stoichiometry. Those results indicating a length- and functional-dependence of operon stoichiometry control for larger operons. The RNA–protein correlation also echoed this trend (**Figure 4B**). In both strains, the Pearson R^2 of the RNA–protein correlation in "Complex" group is considerably higher than in "Pathway" group (**Figure 4C**). These results indicated that there may be some inherent difference between "Pathway" and "Complex" type operons.

Significant difference ($p < 0.01$) of protein CV distribution were observed among the 2-protein operons against larger (4-/5-protein) operons only in "Complex" subgroups, but not observed in the "Pathway" operon subgroups in both BW25113 and MG1655 strains (**Figure 4D**, see details in **Supplementary Table S4**), consolidated our abovementioned observation that the shorter operons among "Complex" groups operons tend to be regulated more stringently.

Enzyme Activity Correlates to the Differential Translation of "Pathway" Genes

Figure 5A showed examples of "Complex" and "Pathway" operons, respectively. Genes in three operons showed almost same RNA abundance. However, the operon ID 3641 encoding ribosome proteins showed similar protein abundance, while the operon ID 3767 encoding enzymes in arginine synthesis pathway and the operon ID 3157 encoding enzymes in biotin synthesis pathway showed exaggerated difference in protein abundance. Both translation and degradation may affect the protein abundance. We found no significant difference of the "Complex" and "Pathway" operons in terms of the CV of protein half-life within operons (KS-test, $P = 0.932$) (**Supplementary Figure S5**). Therefore, the translational regulation should be the major factor leading to such differential protein expression in "Pathway" operons.

It is expected that the "Complex" proteins tend to be tightly stoichiometrically controlled to build functional complexes. In contrast, what benefit is related to the differential translation of the "Pathway" genes in one operon?

Since "Pathway" proteins comprise pathways, they sequentially catalyze conversion of a substrate to product via multiple and successive reaction steps. Therefore, we hypothesize that the bacteria require less high-efficiency enzymes to avoid energy waste. **Supplementary Figure S6** showed the arginine synthesis pathway in *E. coli*, where ArgB and ArgC are two enzymes that relay. The Michalis constant (K_M) of ArgB is more than 3 times higher than ArgC (**Table 1**), showing that the binding of ArgB to the substrate is weaker. Although lacking the measured value of ArgB k_{cat} in *E. coli*, ArgB k_{cat} value in yeast (56% homology in amino acid sequence) is approximately 1/3 of ArgC in *E. coli*, as a reference. These data indicated that the catalytic efficiency of ArgC is higher than ArgB. Therefore, ArgC is less needed in *E. coli*, which matched the proteome quantification (**Figure 5A**). No argH activity data is available in *E. coli*. The homologous enzyme in *Anas platyrhynchos* (75% homology in amino acid sequence) showed also lower catalytic efficiency than ArgC. Similar trend was also observed in the biotin synthesis pathway (**Supplementary Figure S6**). The enzyme BioF possess k_{cat} value one order of magnitude higher than BioB, indicating higher conversion efficiency. Therefore, BioF is much less produced in *E. coli*. The enzyme BioC is extremely efficient with three orders of magnitude higher k_{cat} value than BioF and much lower K_M value than bioB and BioF. Therefore, BioC is orders of magnitude lower than the other two enzymes in *E. coli*.

To further validate this hypothesis, we manually went through all quantified "Pathway" operons and found other nine operons which fit the following criteria: (a) at least two quantified proteins in one operons, and they must exist in the same metabolic pathway; (b) their pathway must be branch-free, i.e., the substrate should be converted sequentially by these enzymes without introducing other rate-limiting metabolites as branches in pathway, at least in the range of the quantified enzymes (illustrations see **Figure 5B**); (c) enzyme activity parameters, e.g., k_{cat} or K_M, should be available, and the activity parameters of at least one enzyme should be available in *E. coli*. The RNA abundance, protein abundance and the K_M values are illustrated in **Figure 5C**. All of the nine operons validated the correlation of the enzyme activity and the protein abundance without exception: less active enzymes (represented by higher K_M values) were expressed in higher amount at protein level. To be

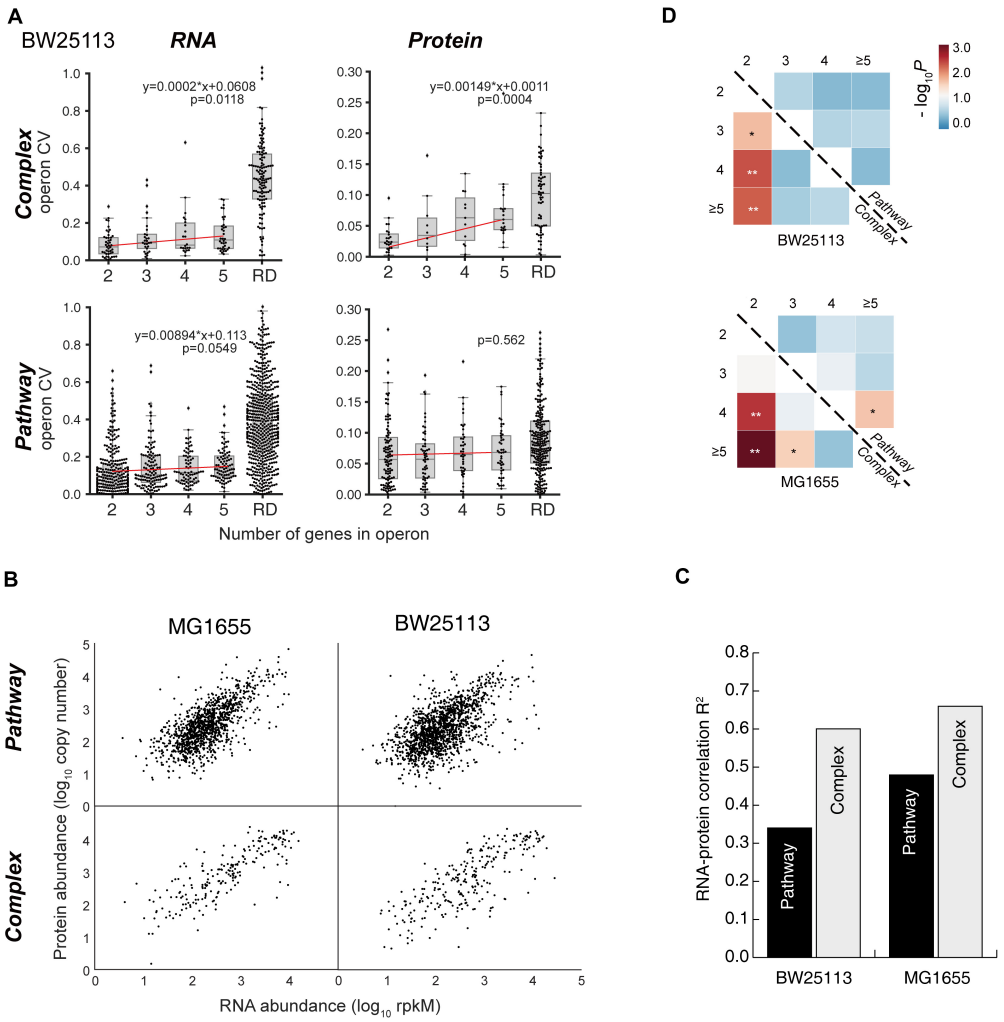

FIGURE 4 | Length-dependence of operon stoichiometry control. **(A)** Linear regression analysis of gene CVs at RNA and protein levels within operon that encodes proteins forming complexes or involving in the metabolic pathways in BW25113 strain. *P*-values of the regression are indicated in the plots. $P < 0.05$ are considered significant. RD, randomized control. **(B)** The RNA–protein scatter plots of the gene expression levels of the "Complex" and "Pathway" operons in two strains, respectively. **(C)** The correlation coefficient R^2 of the RNA–protein correlation shown in **(B)** panel. **(D)** The mutual *P*-value (Mann–Whitney *U* test) matrix of the protein CV within "Pathway" operons and "Complex" operons, respectively. $*P < 0.05$; $**P < 0.01$.

noted, six out of nine operons showed inverse proportion of RNA and protein expression, suggesting that such regulation is mainly conducted at translation level.

Divergence of Stoichiometry Control Is Regulated at Translation Level by the Gene Intervals

Next, we investigated the factor that could determine the lower stringency of "Pathway" groups of operons. As both groups of operons contain a wide variety of genes, the major difference should lay on the gene structures. This is reflected by the gene intervals, defined as the distance from the stop codon of the first gene to the start codon of the next gene downstream within one operon (**Figure 6A**). A strongly significant difference on gene intervals was observed between two groups on their gene interval distributions ($p = 1.314 \times 10^{-5}$, KS-test, **Figure 6B**). The genes in "Complex" genes tend to be arranged very near to each other, while the "Pathway" genes tend to be located far from each other, reflected by the larger mean and median value of the gene interval.

Both a 70S ribosome and a 30S subunit cover about 40 nucleotides of the mRNA, roughly 18–20 nucleotides upstream and 16–18 nucleotides downstream of the P-site codon (Beyer et al., 1994). If a terminating 70S ribosome would have a downstream initiation site nearer than 40 nucleotides, a 70S termination event and a downstream 30S-binding initiation could not be independent events due to a steric clash between 70S and 30S. In this case, the 70S scanning initiation mode plays a major role for translation, meaning the first cistron is initiated by 30S and the downstream cistrons by 70S

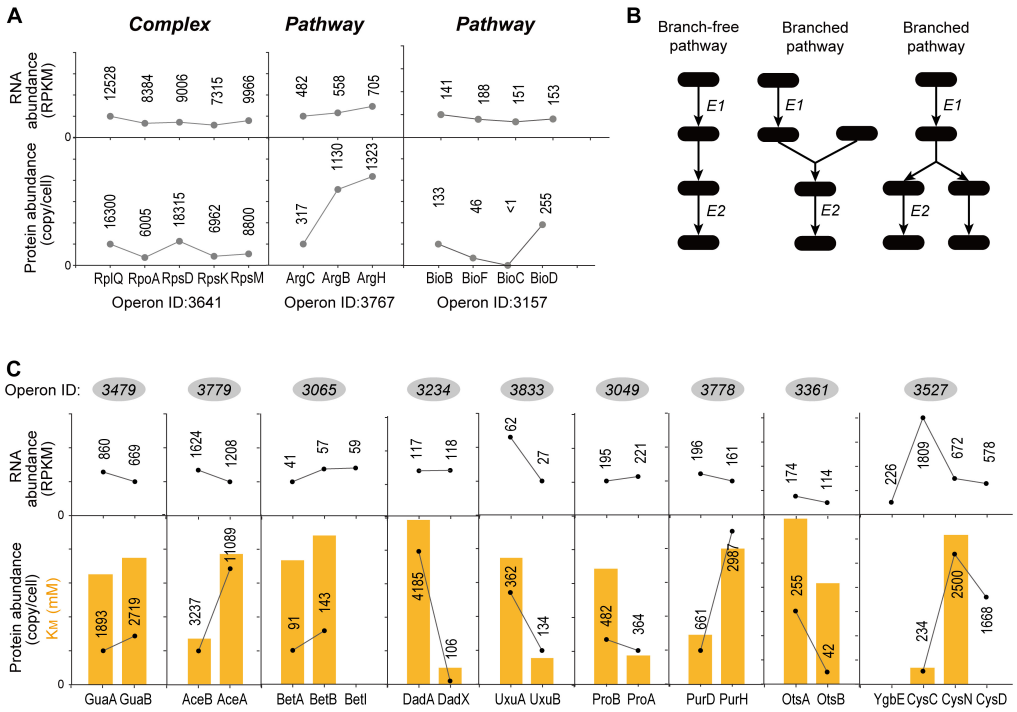

FIGURE 5 | Examples of "Complex" and "Pathway" operons. **(A)** The gene expression at RNA and protein levels in *E. coli* K-12 MG1655 strain. The detailed expression level are marked on the data points. <1 means the expression level is too low to be confidently quantified. **(B)** Illustrations of the "branch-free" and "branched" pathways. E1 and E2 represent the quantified proteins in the same operon. **(C)** All nine operons which encodes enzymes that exist in branch-free pathways. Their RNA abundance, protein abundance and K_M values were plotted. K_M values are plotted in orange bars. Detailed K_M values are listed in **Table 1**. Detailed pathways are illustrated in **Supplementary Figure S6**.

scanning to achieve a strict and precise 1:1 stoichiometric ratio (Yamamoto et al., 2016). This may be one of the factors explained that the "Complex" operons, with shorter gene intervals, are more stringently regulated for stoichiometry.

DISCUSSION

Near-complete coverage of proteome identification and quantification is always a goal of proteomics research, as it reveals detailed global dynamics of important biological processes; for example, the dormancy and resuscitation of *Mycobacterium tuberculosis* (Schubert et al., 2015). The high proteomic resolution of quantification allows in-depth investigation of many scientific questions in debate for decades. In this study, we employed DIA based HRM-MS, a highly reproducible label-free quantification method in *E. coli* strains BW25113 and MG1655. Our datasets were better than the previously reported DDA-based *E. coli* proteome quantification datasets in terms of reproducibility. We therefore generated the most complete investigation on abundance of proteins encoded within operons in *E. coli* based on tryptic digestion up to date, which allows us to accurately evaluate the stoichiometry presumption of the operon organization. To further increase the sensitivity of identification, other complement approaches such as LysC digestion might be helpful (Wisniewski and Rakus, 2014). Nevertheless, since we have already quantified 93.6% of the cytosolic proteins, specific methods dealing with the hydrophobic nature of membrane proteins should be employed to further expand the proteome coverage. The DIA based HRM-MS relies heavily on DDA spectral libraries. Therefore, they share the same shortcomings, e.g., the dependence on the physical and chemical properties of proteins.

We confirmed from our HRM-MS results that the operons coordinate the gene expression more stringent than the randomized control in general. In addition, we found a multifaceted nature of the operon regulation: operons are not created equal. The stringency is length-dependent and functional-dependent at protein level. Such multifaceted organization of operons revealed two-level control: the operon unified transcriptional level and gene-specific translational level, which benefits the host in different aspects.

Although the operon organization maintains in general the stoichiometry of the genes in the operons compared to fully randomized scenario, the operons for metabolic pathways are in general less strictly controlled for stoichiometry balance compared to those operons for protein complexes. Protein complexes needs stoichiometry to maintain their functions. Therefore, the operons encoding protein complexes are tightly regulated to ensure the equal expression, such as the ribosome protein operons (**Figure 5A**). In contrast, since the operons for metabolic pathways are not necessarily

TABLE 1 | Kinetic parameters of the enzymes in one metabolic pathway.

Enzyme	k_{cat} (s^{-1})	K_M (mM)	Homology to *E. coli* protein (for non-*E. coli* proteins)	References
ArgB	4.9 (*S. cerevisiae*)	1.3	56%	Gil-Ortiz et al., 2010; de Cima et al., 2012
ArgC	14	0.4		McLoughlin and Copley, 2008
ArgH	4.9 (*Anas platyrhynchos*)	0.4 (*Anas platyrhynchos*)	75%	Chakraborty et al., 1999
BioB	0.0039	0.002		Farh et al., 2001; Taylor et al., 2008
BioF	0.05825	0.025		Turbeville et al., 2011
BioC	98.334 (*Bacillus cereus*)	0.00108 (*Bacillus cereus*)	62%	Lin and Cronan, 2012
GuaB	13	0.061		Kerr and Hedstrom, 1997
GuaA	23	0.053		Oliver et al., 2014
AceB		0.022		Lohman et al., 2008
AceA		0.063		MacKintosh and Nimmo, 1988
BetA		1.5		Landfald and Strom, 1986
BetB		1.8		Gruez et al., 2004
UxuA		4.79		Qiu et al., 2012
UxuB		1		Hickman and Ashwell, 1960
DadA		30		Franklin and Venables, 1976
DadX		3.03		Wu et al., 2008
ProB		1.2 (*Pseudomonas aeruginosa*)	74%	Krishna and Leisinger, 1979
ProA		300		McLoughlin and Copley, 2008
OtsA		1 (*Thermoplasma acidophilum*)	65%	Gao et al., 2014
OtsB		0.61		Kuznetsova et al., 2006
CysC		0.0005		Satishchandran and Markham, 2000
CysD + CysN		0.0045 (*Thiobacillus denitrificans*)	67% (CysD)	Gay et al., 2009
PurD		0.03		Cheng et al., 1990
PurH		0.082 (*Methanocaldococcus jannaschii*)	50%	Graupner et al., 2002

Values are measured in E. coli unless specified. Non-E. coli data were used as a reference when no E. coli data is available.

forming a complex for their functions, their distinct specific activities set their specific demand in quantity. All quantified operons which encode enzymes in branch-free pathways and with available enzyme activity data validated this hypothesis without exception (**Figure 5C**). Nevertheless, bacteria need to regulate related metabolic pathways in quick response to environmental stimuli. Therefore, organizing the proteins in the same pathway under the control of one promoter would minimize the regulatory complexity of the adjustments, leaving the delicate control of each individual gene to the translational level. The available data indicated that such translational regulation is quite common (**Figure 5C**).

Our data also showed that the shorter operons, whose products form complexes, tend to be more tightly controlled in stoichiometric expression (**Figure 4**). This could be understandable such as two-component protein complexes would be invalid if there were imbalanced expression of the components. For instance, the assembly efficiency decreased remarkably if two subunits of bacterial luciferase *LuxA* and *LuxB* were split at distant chromosomal sites (Shieh et al., 2015). Large operons tend to encode proteins for large complexes such as ribosomes (**Figure 4**). Such large complexes would sustain for relatively long time in cells to perform essential functions. For example, half-life of mammalian ribosomes can be as long as 300 h (Nikolov et al., 1983). Many protein components of ribosomes are dissociable and interchangeable with unbound counterparts (Schafer et al., 2003). This allows the rapid renewal of the damaged proteins of the complex without degrading and re-synthesizing the entire complex, which is the most energy-efficient way to keep these valuable complexes in good condition. This requires the delicate expression control of these proteins within one operon to meet the actual demand.

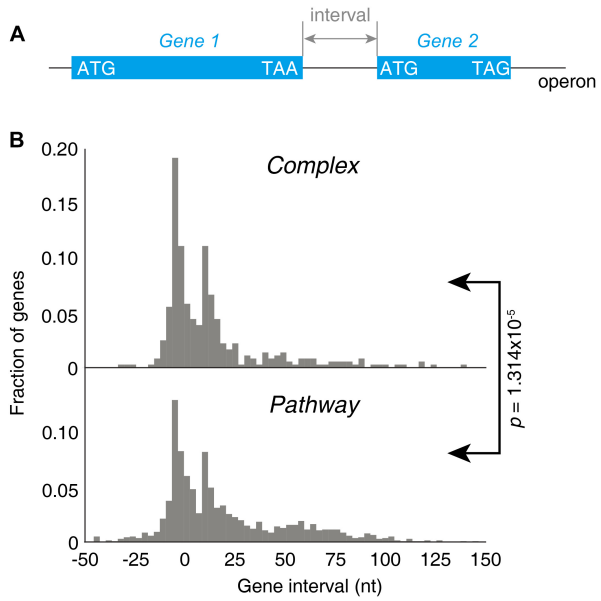

FIGURE 6 | Gene intervals in the "Complex" and "Pathway" operons. **(A)** Illustration of the gene interval within one operon. **(B)** Distribution of the gene intervals of "Complex" and "Pathway" operons.

In another aspect, the differential expression regulation within an operon is also important for bacteria. Previous studies proposed that such regulation happens via generating different transcripts from multiple promotors/terminators [e.g., *Bacillus subtilis dnaK* operon (Homuth et al., 1999), *Vibrio vulnificus putAP* operon (Lee et al., 2003), *Zymomonas mobilis gap* operon (Eddy et al., 1989), *E. coli glpEGR* operon (Yang and Larson, 1998)], or via differential degrading mRNAs [e.g., *Acinetobacter calcoaceticus mop* operon (Schirmer and Hillen, 1998), *E. coli atp* operon (McCarthy, 1990; McCarthy et al., 1991)]. However, these studies included only individual cases of specific operons. Taking advantage of deep coverage of *E. coli* cytosolic proteome, our data indicated that enzyme activity seems to be an additional driving force for the differential expression regulation within an operon. Highly efficient enzymes tended to be less produced than the other counterparts in the same pathway. In such cases, deviating from stoichiometry minimizes the energy waste and thus may provide survival advantage. This explained the fact that no significant length-dependent stoichiometry is observed in "Pathway" proteins. Our analysis was restricted by the very limited availability of the enzyme activity and kinetics data. Validation using more such data is necessary in the future.

In this study, we noticed that the gene intervals in operons may serve as a regulatory factor for stoichiometry. It is a general accepted notion that termination of bacterial protein synthesis is obligatorily followed by recycling step governed by the factors ribosomal recycling factor (RRF), EF-G, and IF3, where the ribosome dissociates into its subunits (Hirokawa et al., 2006). In contrast, a recently described 70S-scanning mode of initiation holds that after termination, the 70S ribosomes do not dissociate after termination step but rather scan along with the mRNA until reaching the initiation site of the downstream cistron of the same mRNA (Yamamoto et al., 2016). Binding of fMet-tRNA triggers 70S scanning, which occurs in the absence of energy-rich compounds (e.g., ATP, GTP) and seems to be driven by unidimensional diffusion (Yamamoto et al., 2016). Therefore, the 70S scanning initiation might be a mechanism to read out the "stoichiometry code" of closely located genes in operons. In addition, the rate of translation of an ORF is controlled by a number of other mRNA features. For example, the codon selection and the cognate tRNA concentration dictate the translational pausing, which is a strong determinant of co-translational folding for most proteins (Zhang et al., 2009; Zhong et al., 2015; Lian et al., 2016). Shine–Dalgarno (SD) sequence accessibility and strength have been implicated in translational initiation (Steitz and Jakes, 1975). Genome-wide mRNA secondary structure analysis indicated that ORF translation rate is correlated with its mRNA structure in bacteria (Burkhardt et al., 2017), but not in mammalian cells (Lian et al., 2016). Although highly stable mRNA structures in direct proximity to the initiation codon diminish translation efficiency (de Smit and van Duin, 1990), secondary structure hiding the SD sequence in front of the second cistron prevents 30S binding initiation, but the secondary structure can be resolved by scanning 70S ribosomes when the secondary structure has a comparable stability ($\Delta G \geq -6$ kcal/mol at 30°C) (Qin et al., 2016; Yamamoto et al., 2016). These molecular mechanisms should be universal beyond exponential growth condition, although further validation would be needed.

This two-level regulation mode involving transcription and translation would balance the regulation in different time-scale. As transcriptional control takes effects at least in half an hour, it is suitable for sustained alteration in gene expression and in pathway-scale. In contrast, rapid and fine adjustment can be only performed at the translational level, which takes effects in less than 1 min and occurs at the individual gene level (Zhong et al., 2015). This rapid responsiveness would be also ideal for real-time adjustment of the proteins needed in complexes and pathways. Translational regulation largely contributes to the proteome complexity and minimizes the energy waste on synthesizing unnecessary proteins. Thus, the delicate and differential translational regulation in bacteria maintains the functionality and efficiency of both macromolecular complexes and metabolic pathways, which is a desperate need of the survival and competence of bacteria.

DATA AVAILABILITY

The datasets generated for this study can be found in GEO, GSM3489376, GSM3489377.

AUTHOR CONTRIBUTIONS

GZ, JZ, and CS: conceptualization. JZ and HZ: experiment and visualization. JZ, HZ, BQ, and RN: data analysis. GZ and Q-YH: supervision. JZ, HZ, BQ, RN, Q-YH, CS, and GZ: writing. GZ: funding acquisition.

FUNDING

This work was supported by the Ministry of Science and Technology of China, National Key Research and Development Program (2017YFA0505001/2017YFA0505101/2018YFC0910201/2018YFC0910202), Guangdong Key R&D Program (2019B020226001), and the Distinguished Young Talent Award of National High-level Personnel Program of China.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.00473/full#supplementary-material

FIGURE S1 | Physical and chemical features of quantified and unquantified proteins in *E. coli*. Panels **(A–D)** are distributions of protein isoelectric point, instability, protein length, and hydrophobicity. Instability tests a protein for stability, which value above 40 means the protein has a shorter half-life. The Kolmogorov–Smirnov test (KS-test) was used to test the distribution between quantified and unquantified proteins. Significance different were found between quantified and unquantified proteins indicated different physical–chemical characterizations of them.

FIGURE S2 | CV within operons of measured data and randomized negative control at RNA level.

FIGURE S3 | Length- and functional-dependence of operon stoichiometry control compare to randomized data. **(A–E)** Distribution among "Complex," "Complex_random," "Pathway," and "Pathway_random" groups in BW25113. **(F–J)** Distribution among "Complex," "Complex_random," "Pathway," "Pathway_random" groups in MG1655. In most cases, the CV in "Complex" and "Pathway" groups were lower than corresponding "Complex_random" and "Pathway_random" groups in both strains, indicating real robust signals but not statistical artifacts in operons present in *E. coli*.

FIGURE S4 | Length dependence of the stoichiometry control at RNA and protein levels. Same figure in MG1655 strain, comparable to the **Figure 4A**.

FIGURE S5 | Distribution of protein half-life CV in "Complex" and "Pathway" operons.

FIGURE S6 | Pathways in **Figure 5**.

TABLE S1 | DIA MS/MS isolation windows table.

TABLE S2 | Spectral library used in HRM-MS analysis.

TABLE S3 | Protein abundances and properties used in this analysis.

TABLE S4 | *P*-value of Mann–Whitney *U* test between complex and pathway operons shown in **Figure 3C**.

TABLE S5 | *P*-value of gene ontology (GO) overrepresentation of the operons with lower and higher CV groups shown in **Figures 3D,E**.

SUPPLEMENTARY SCRIPTS | The python scripts used in this study.

SUPPLEMENTARY DATA | The relationship of submitted raw data.

REFERENCES

Arike, L., Valgepea, K., Peil, L., Nahku, R., Adamberg, K., and Vilu, R. (2012). Comparison and applications of label-free absolute proteome quantification methods on *Escherichia coli*. *J. Proteom.* 75, 5437–5448. doi: 10.1016/j.jprot.2012.06.020

Bartholomaus, A., Fedyunin, I., Feist, P., Sin, C., Zhang, G., Valleriani, A., et al. (2016). Bacteria differently regulate mRNA abundance to specifically respond to various stresses. *Philos. Trans. A Math. Phys. Eng. Sci.* 374:20150069. doi: 10.1098/rsta.2015.0069

Ben-Shahar, Y., Nannapaneni, K., Casavant, T. L., Scheetz, T. E., and Welsh, M. J. (2007). Eukaryotic operon-like transcription of functionally related genes in Drosophila. *Proc. Natl. Acad. Sci. U.S.A.* 104, 222–227. doi: 10.1073/pnas.0609683104

Beyer, D., Skripkin, E., Wadzack, J., and Nierhaus, K. H. (1994). How the ribosome moves along the mRNA during protein synthesis. *J. Biol. Chem.* 269, 30713–30717.

Bloom, J. S., Khan, Z., Kruglyak, L., Singh, M., and Caudy, A. A. (2009). Measuring differential gene expression by short read sequencing: quantitative comparison to 2-channel gene expression microarrays. *BMC Genom.* 10:221. doi: 10.1186/1471-2164-10-221

Blumenthal, T. (2004). Operons in eukaryotes. *Brief. Funct. Genomic Proteom.* 3, 199–211. doi: 10.1093/bfgp/3.3.199

Burkhardt, D. H., Rouskin, S., Zhang, Y., Li, G. W., Weissman, J. S., and Gross, C. A. (2017). Operon mRNAs are organized into ORF-centric structures that predict translation efficiency. *eLife* 6:e22037. doi: 10.7554/eLife.22037

Chakraborty, A. R., Davidson, A., and Howell, P. L. (1999). Mutational analysis of amino acid residues involved in argininosuccinate lyase activity in duck delta II crystallin. *Biochemistry* 38, 2435–2443. doi: 10.1021/bi982150g

Chapman, J. D., Goodlett, D. R., and Masselon, C. D. (2014). Multiplexed and data-independent tandem mass spectrometry for global proteome profiling. *Mass Spectrom. Rev.* 33, 452–470. doi: 10.1002/mas.21400

Cheng, Y. S., Shen, Y., Rudolph, J., Stern, M., Stubbe, J., Flannigan, K. A., et al. (1990). Glycinamide ribonucleotide synthetase from *Escherichia coli*: cloning, overproduction, sequencing, isolation, and characterization. *Biochemistry* 29, 218–227. doi: 10.1021/bi00453a030

de Cima, S., Gil-Ortiz, F., Crabeel, M., Fita, I., and Rubio, V. (2012). Insight on an arginine synthesis metabolon from the tetrameric structure of yeast acetylglutamate kinase. *PLoS One* 7:e34734. doi: 10.1371/journal.pone.0034734

de Smit, M. H., and van Duin, J. (1990). Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. *Proc. Natl. Acad. Sci. U.S.A.* 87, 7668–7672. doi: 10.1073/pnas.87.19.7668

Doerr, A. (2015). DIA mass spectrometry. *Nat. Methods* 12, 35–35. doi: 10.1038/nmeth.3234

Eddy, C. K., Mejia, J. P., Conway, T., and Ingram, L. O. (1989). Differential expression of gap and pgk genes within the gap operon of *Zymomonas mobilis*. *J. Bacteriol.* 171, 6549–6554. doi: 10.1128/jb.171.12.6549-6554.1989

Farh, L., Hwang, S. Y., Steinrauf, L., Chiang, H. J., and Shiuan, D. (2001). Structure-function studies of *Escherichia coli* biotin synthase via a chemical modification and site-directed mutagenesis approach. *J. Biochem.* 130, 627–635. doi: 10.1093/oxfordjournals.jbchem.a003028

Franklin, F. C., and Venables, W. A. (1976). Biochemical, genetic, and regulatory studies of alanine catabolism in *Escherichia coli* K12. *Mol. Gen. Genet.* 149, 229–237. doi: 10.1007/bf00332894

Gao, Y., Jiang, Y., Liu, Q., Wang, R., Liu, X., and Liu, B. (2014). Enzymatic and regulatory properties of the trehalose-6-phosphate synthase from the thermoacidophilic archaeon *Thermoplasma acidophilum*. *Biochimie* 101, 215–220. doi: 10.1016/j.biochi.2014.01.018

Gay, S. C., Fribourgh, J. L., Donohoue, P. D., Segel, I. H., and Fisher, A. J. (2009). Kinetic properties of ATP sulfurylase and APS kinase from *Thiobacillus denitrificans*. *Arch. Biochem. Biophys.* 489, 110–117. doi: 10.1016/j.abb.2009.07.026

Gil-Ortiz, F., Ramon-Maiques, S., Fernandez-Murga, M. L., Fita, I., and Rubio, V. (2010). Two crystal structures of *Escherichia coli* N-acetyl-L-glutamate kinase demonstrate the cycling between open and closed conformations. *J. Mol. Biol.* 399, 476–490. doi: 10.1016/j.jmb.2010.04.025

Gordon, S. P., Tseng, E., Salamov, A., Zhang, J., Meng, X., Zhao, Z., et al. (2015). Widespread polycistronic transcripts in fungi revealed by single-molecule mRNA sequencing. *PLoS One* 10:e0132628. doi: 10.1371/journal.pone.0132628

Graupner, M., Xu, H., and White, R. H. (2002). New class of IMP cyclohydrolases in Methanococcus jannaschii. *J. Bacteriol.* 184, 1471–1473. doi: 10.1128/jb.184.5.1471-1473.2002

Gruez, A., Roig-Zamboni, V., Grisel, S., Salomoni, A., Valencia, C., Campanacci, V., et al. (2004). Crystal structure and kinetics identify *Escherichia coli* YdcW gene product as a medium-chain aldehyde dehydrogenase. *J. Mol. Biol.* 343, 29–41. doi: 10.1016/j.jmb.2004.08.030

Guell, M., van Noort, V., Yus, E., Chen, W. H., Leigh-Bell, J., Michalodimitrakis, K., et al. (2009). Transcriptome complexity in a genome-reduced bacterium. *Science* 326, 1268–1271. doi: 10.1126/science.1176951

Haft, R. J., Keating, D. H., Schwaegler, T., Schwalbach, M. S., Vinokur, J., Tremaine, M., et al. (2014). Correcting direct effects of ethanol on translation and transcription machinery confers ethanol tolerance in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 111, E2576–E2585. doi: 10.1073/pnas.1401853111

Hickman, J., and Ashwell, G. (1960). Uronic acid metabolism in bacteria. II. Purification and properties of D-altronic acid and D-mannonic acid dehydrogenases in *Escherichia coli*. *J. Biol. Chem.* 235, 1566–1570.

Hirokawa, G., Demeshkina, N., Iwakura, N., Kaji, H., and Kaji, A. (2006). The ribosome-recycling step: consensus or controversy?. *Trends Biochem. Sci.* 31, 143–149. doi: 10.1016/j.tibs.2006.01.007

Homuth, G., Mogk, A., and Schumann, W. (1999). Post-transcriptional regulation of the Bacillus subtilis dnaK operon. *Mol. Microbiol.* 32, 1183–1197. doi: 10.1046/j.1365-2958.1999.01428.x

Huang, D. W., Sherman, B. T., Tan, Q., Kir, J., Liu, D., Bryant, D., et al. (2007). DAVID bioinformatics resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res.* 35, W169–W175. doi: 10.1093/nar/gkm415

Iber, D. (2006). A quantitative study of the benefits of co-regulation using the spoIIA operon as an example. *Mol. Syst. Biol.* 2:43. doi: 10.1038/msb4100084

Ishihama, Y., Schmidt, T., Rappsilber, J., Mann, M., Hartl, F. U., Kerner, M. J., et al. (2008). Protein abundance profiling of the *Escherichia coli* cytosol. *BMC Genom.* 9:102. doi: 10.1186/1471-2164-9-102

Kerr, K. M., and Hedstrom, L. (1997). The roles of conserved carboxylate residues in IMP dehydrogenase and identification of a transition state analog. *Biochemistry* 36, 13365–13373. doi: 10.1021/bi9714161

Krishna, R. V., and Leisinger, T. (1979). Biosynthesis of proline in *Pseudomonas aeruginosa*. Partial purification and characterization of gamma-glutamyl kinase. *Biochem. J.* 181, 215–222. doi: 10.1042/bj1810215

Kuznetsova, E., Proudfoot, M., Gonzalez, C. F., Brown, G., Omelchenko, M. V., Borozan, I., et al. (2006). Genome-wide analysis of substrate specificities of the *Escherichia coli* haloacid dehalogenase-like phosphatase family. *J. Biol. Chem.* 281, 36149–36161. doi: 10.1074/jbc.M605449200

Landfald, B., and Strom, A. R. (1986). Choline-glycine betaine pathway confers a high level of osmotic tolerance in *Escherichia coli*. *J. Bacteriol.* 165, 849–855. doi: 10.1128/jb.165.3.849-855.1986

Law, K. P., and Lim, Y. P. (2013). Recent advances in mass spectrometry: data independent analysis and hyper reaction monitoring. *Expert Rev. Proteom.* 10, 551–566. doi: 10.1586/14789450.2013.858022

Lawrence, J. G., and Roth, J. R. (1996). Selfish operons: horizontal transfer may drive the evolution of gene clusters. *Genetics* 143, 1843–1860.

Lee, J. H., Park, N. Y., Lee, M. H., and Choi, S. H. (2003). Characterization of the Vibrio vulnificus putAP operon, encoding proline dehydrogenase and proline permease, and its differential expression in response to osmotic stress. *J. Bacteriol.* 185, 3842–3852. doi: 10.1128/jb.185.13.3842-3852.2003

Li, G. W., Burkhardt, D., Gross, C., and Weissman, J. S. (2014). Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell* 157, 624–635. doi: 10.1016/j.cell.2014.02.033

Lian, X., Guo, J., Gu, W., Cui, Y., Zhong, J., Jin, J., et al. (2016). Genome-wide and experimental resolution of relative translation elongation speed at individual gene level in human cells. *PLoS Genet.* 12:e1005901. doi: 10.1371/journal.pgen.1005901

Lin, S., and Cronan, J. E. (2012). The BioC O-methyltransferase catalyzes methyl esterification of malonyl-acyl carrier protein, an essential step in biotin synthesis. *J. Biol. Chem.* 287, 37010–37020. doi: 10.1074/jbc.M112.410290

Liu, W., Xiang, L., Zheng, T., Jin, J., and Zhang, G. (2017). TranslatomeDB: a comprehensive database and cloud-based analysis platform for translatome sequencing data. *Nucleic Acids Res.* 46, D206–D212. doi: 10.1093/nar/gkx1034

Lohman, J. R., Olson, A. C., and Remington, S. J. (2008). Atomic resolution structures of *Escherichia coli* and Bacillus anthracis malate synthase A: comparison with isoform G and implications for structure-based drug discovery. *Protein Sci.* 17, 1935–1945. doi: 10.1110/ps.036269.108

MacKintosh, C., and Nimmo, H. G. (1988). Purification and regulatory properties of isocitrate lyase from *Escherichia coli* ML308. *Biochem. J.* 250, 25–31. doi: 10.1042/bj2500025

Maier, T., Schmidt, A., Guell, M., Kuhner, S., Gavin, A. C., Aebersold, R., et al. (2011). Quantification of mRNA and protein and integration with protein turnover in a bacterium. *Mol. Syst. Biol.* 7:511. doi: 10.1038/msb.2011.38

Mao, X., Ma, Q., Zhou, C., Chen, X., Zhang, H., Yang, J., et al. (2014). DOOR 2.0: presenting operons and their functions through dynamic and integrated views. *Nucleic Acids Res.* 42, D654–D659. doi: 10.1093/nar/gkt1048

McCarthy, J. E. (1990). Post-transcriptional control in the polycistronic operon environment: studies of the atp operon of *Escherichia coli*. *Mol. Microbiol.* 4, 1233–1240. doi: 10.1111/j.1365-2958.1990.tb00702.x

McCarthy, J. E., Gerstel, B., Surin, B., Wiedemann, U., and Ziemke, P. (1991). Differential gene expression from the *Escherichia coli* atp operon mediated by segmental differences in mRNA stability. *Mol. Microbiol.* 5, 2447–2458. doi: 10.1111/j.1365-2958.1991.tb02090.x

McLoughlin, S. Y., and Copley, S. D. (2008). A compromise required by gene sharing enables survival: implications for evolution of new enzyme activities. *Proc. Natl. Acad. Sci. U.S.A.* 105, 13497–13502. doi: 10.1073/pnas.0804804105

Mi, H., Huang, X., Muruganujan, A., Tang, H., Mills, C., Kang, D., et al. (2017). PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res.* 45, D183–D189. doi: 10.1093/nar/gkw1138

Neilson, K. A., Ali, N. A., Muralidharan, S., Mirzaei, M., Mariani, M., Assadourian, G., et al. (2011). Less label, more free: approaches in label-free quantitative mass spectrometry. *Proteomics* 11, 535–553. doi: 10.1002/pmic.201000553

Nikolov, E. N., Dabeva, M. D., and Nikolov, T. K. (1983). Turnover of ribosomes in regenerating rat liver. *Int. J. Biochem.* 15, 1255–1260. doi: 10.1016/0020-711x(83)90215-x

Oliver, J. C., Gudihal, R., Burgner, J. W., Pedley, A. M., Zwierko, A. T., Davisson, V. J., et al. (2014). Conformational changes involving ammonia tunnel formation and allosteric control in GMP synthetase. *Arch. Biochem. Biophys.* 545, 22–32. doi: 10.1016/j.abb.2014.01.004

Osbourn, A. E., and Field, B. (2009). Operons. *Cell. Mol. Life Sci.* 66, 3755–3775. doi: 10.1007/s00018-009-0114-3

Pi, H., Lee, L. W., and Lo, S. J. (2009). New insights into polycistronic transcripts in eukaryotes. *Chang Gung Med. J.* 32, 494–498.

Price, M. N., Huang, K. H., Alm, E. J., and Arkin, A. P. (2005). A novel method for accurate operon predictions in all sequenced prokaryotes. *Nucleic Acids Res.* 33, 880–892. doi: 10.1093/nar/gki232

Purvine, S., Eppel, J. T., Yi, E. C., and Goodlett, D. R. (2003). Shotgun collision-induced dissociation of peptides using a time of flight mass analyzer. *Proteomics* 3, 847–850. doi: 10.1002/pmic.200300362

Qin, B., Yamamoto, H., Ueda, T., Varshney, U., and Nierhaus, K. H. (2016). The termination phase in protein synthesis is not obligatorily followed by the RRF/EF-G-dependent recycling phase. *J. Mol. Biol.* 428, 3577–3587. doi: 10.1016/j.jmb.2016.05.019

Qiu, X., Tao, Y., Zhu, Y., Yuan, Y., Zhang, Y., Liu, H., et al. (2012). Structural insights into decreased enzymatic activity induced by an insert sequence in mannonate dehydratase from Gram negative bacterium. *J. Struct. Biol.* 180, 327–334. doi: 10.1016/j.jsb.2012.06.013

Satishchandran, C., and Markham, G. D. (2000). Mechanistic studies of *Escherichia coli* adenosine-5′-phosphosulfate kinase. *Arch. Biochem. Biophys.* 378, 210–215. doi: 10.1006/abbi.2000.1841

Schafer, T., Strauss, D., Petfalski, E., Tollervey, D., and Hurt, E. (2003). The path from nucleolar 90S to cytoplasmic 40S pre-ribosomes. *EMBO J.* 22, 1370–1380. doi: 10.1093/emboj/cdg121

Schirmer, F., and Hillen, W. (1998). The Acinetobacter calcoaceticus NCIB8250 mop operon mRNA is differentially degraded, resulting in a higher level of the 3' CatA-encoding segment than of the 5′ phenolhydroxylase-encoding portion. *Mol. Gen. Genet.* 257, 330–337. doi: 10.1007/pl00008621

Schmidt, A., Beck, M., Malmstrom, J., Lam, H., Claassen, M., Campbell, D., et al. (2011). Absolute quantification of microbial proteomes at different states by directed mass spectrometry. *Mol. Syst. Biol.* 7:510. doi: 10.1038/msb.2011.37

Schubert, O. T., Ludwig, C., Kogadeeva, M., Zimmermann, M., Rosenberger, G., Gengenbacher, M., et al. (2015). Absolute Proteome composition and dynamics during dormancy and resuscitation of *Mycobacterium tuberculosis*. *Cell Host Microbe* 18, 96–108. doi: 10.1016/j.chom.2015.06.001

Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., et al. (2011). Global quantification of mammalian gene expression control. *Nature* 473, 337–342. doi: 10.1038/nature10098

Seshasayee, A. S., Bertone, P., Fraser, G. M., and Luscombe, N. M. (2006). Transcriptional regulatory networks in bacteria: from input signals to output responses. *Curr. Opin. Microbiol.* 9, 511–519. doi: 10.1016/j.mib.2006.08.007

Shieh, Y. W., Minguez, P., Bork, P., Auburger, J. J., Guilbride, D. L., Kramer, G., et al. (2015). Operon structure and cotranslational subunit association direct protein assembly in bacteria. *Science* 350, 678–680. doi: 10.1126/science.aac8171

Steitz, J. A., and Jakes, K. (1975). How ribosomes select initiator regions in mRNA: base pair formation between the 3' terminus of 16S rRNA and the mRNA during initiation of protein synthesis in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 72, 4734–4738. doi: 10.1073/pnas.72.12.4734

Taylor, A. M., Farrar, C. E., and Jarrett, J. T. (2008). 9-Mercaptodethiobiotin is formed as a competent catalytic intermediate by *Escherichia coli* biotin synthase. *Biochemistry* 47, 9309–9317. doi: 10.1021/bi801035b

Teichmann, S. A., and Babu, M. M. (2004). Gene regulatory network growth by duplication. *Nat. Genet.* 36, 492–496. doi: 10.1038/ng1340

Turbeville, T. D., Zhang, J., Adams, W. C., Hunter, G. A., and Ferreira, G. C. (2011). Functional asymmetry for the active sites of linked 5-aminolevulinate synthase and 8-amino-7-oxononanoate synthase. *Arch. Biochem. Biophys.* 511, 107–117. doi: 10.1016/j.abb.2011.05.002

Wisniewski, J. R., and Rakus, D. (2014). Multi-enzyme digestion FASP and the 'Total Protein Approach'-based absolute quantification of the *Escherichia coli* proteome. *J. Proteom.* 109, 322–331. doi: 10.1016/j.jprot.2014.07.012

Wolf, Y. I., Rogozin, I. B., Kondrashov, A. S., and Koonin, E. V. (2001). Genome alignment, evolution of prokaryotic genome organization, and prediction of gene function using genomic context. *Genome Res.* 11, 356–372. doi: 10.1101/gr.161901

Wu, D., Hu, T., Zhang, L., Chen, J., Du, J., Ding, J., et al. (2008). Residues Asp164 and Glu165 at the substrate entryway function potently in substrate orientation of alanine racemase from *E. coli*: enzymatic characterization with crystal structure analysis. *Protein Sci.* 17, 1066–1076. doi: 10.1110/ps.083495908

Xie, X., and Zubarev, R. A. (2015). Isotopic resonance hypothesis: experimental verification by *Escherichia coli* growth measurements. *Sci. Rep.* 5:9215. doi: 10.1038/srep09215

Yamamoto, H., Wittek, D., Gupta, R., Qin, B., Ueda, T., Krause, R., et al. (2016). 70S-scanning initiation is a novel and frequent initiation mode of ribosomal translation in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 113, E1180–E1189. doi: 10.1073/pnas.1524554113

Yang, B., and Larson, T. J. (1998). Multiple promoters are responsible for transcription of the glpEGR operon of *Escherichia coli* K-12. *Biochim. Biophys. Acta* 1396, 114–126. doi: 10.1016/s0167-4781(97)00179-6

Yun, H., Lee, J. W., Jeong, J., Chung, J., Park, J. M., Myoung, H. N., et al. (2007). EcoProDB: the *Escherichia coli* protein database. *Bioinformatics* 23, 2501–2503. doi: 10.1093/bioinformatics/btm351

Zaslaver, A., Mayo, A. E., Rosenberg, R., Bashkin, P., Sberro, H., Tsalyuk, M., et al. (2004). Just-in-time transcription program in metabolic pathways. *Nat. Genet.* 36, 486–491. doi: 10.1038/ng1348

Zhang, G., Hubalewska, M., and Ignatova, Z. (2009). Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. *Nat. Struct. Mol. Biol.* 16, 274–280. doi: 10.1038/nsmb.1554

Zhang, W., Chen, X., Yan, Z., Chen, Y., Cui, Y., Chen, B., et al. (2017). Detergent-insoluble proteome analysis revealed aberrantly aggregated proteins in human preeclampsia placentas. *J. Proteome Res.* 16, 4468–4480. doi: 10.1021/acs.jproteome.7b00352

Zhao, P., Zhong, J., Liu, W., Zhao, J., and Zhang, G. (2017). Protein-level integration strategy of multiengine MS spectra search results for higher confidence and sequence coverage. *J. Proteome Res.* 16, 4446–4454. doi: 10.1021/acs.jproteome.7b00463

Zheng, Y., Anton, B. P., Roberts, R. J., and Kasif, S. (2005). Phylogenetic detection of conserved gene clusters in microbial genomes. *BMC Bioinform.* 6:243. doi: 10.1186/1471-2105-6-243

Zhong, J., Xiao, C., Gu, W., Du, G., Sun, X., He, Q. Y., et al. (2015). Transfer RNAs mediate the rapid adaptation of *Escherichia coli* to oxidative stress. *PLoS Genet.* 11:e1005302. doi: 10.1371/journal.pgen.1005302

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

REVIEW
published: 04 December 2019
doi: 10.3389/fgene.2019.01153

Coordinated Regulation of Rsd and RMF for Simultaneous Hibernation of Transcription Apparatus and Translation Machinery in Stationary-Phase *Escherichia coli*

Hideji Yoshida[1], Akira Wada[2], Tomohiro Shimada[3,4], Yasushi Maki[1] and Akira Ishihama[4]**

[1] Department of Physics, Osaka Medical College, Takatsuki, Japan, [2] Yoshida Biological Laboratory, Kyoto, Japan, [3] School of Agriculture, Meiji University, Kawasaki, Japan, [4] Research Center for Micro-Nano Technology, Hosei University, Koganei, Japan

OPEN ACCESS

Edited by:
Michael Ibba, The Ohio State University, United States

Reviewed by:
Jonathan Dworkin, Columbia University, United States
Mee-Ngan F. Yap, Northwestern University, United States

***Correspondence:**
Hideji Yoshida
yhide@osaka-med.ac.jp
Akira Ishihama
aishiham@hosei.ac.jp

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *20 July 2019*
Accepted: *22 October 2019*
Published: *04 December 2019*

Citation:
Yoshida H, Wada A, Shimada T, Maki Y and Ishihama A (2019) Coordinated Regulation of Rsd and RMF for Simultaneous Hibernation of Transcription Apparatus and Translation Machinery in Stationary-Phase Escherichia coli. Front. Genet. 10:1153. doi: 10.3389/fgene.2019.01153

Transcription and translation in growing phase of *Escherichia coli*, the best-studied model prokaryote, are coupled and regulated in coordinate fashion. Accordingly, the growth rate-dependent control of the synthesis of RNA polymerase (RNAP) core enzyme (the core component of transcription apparatus) and ribosomes (the core component of translation machinery) is tightly coordinated to keep the relative level of transcription apparatus and translation machinery constant for effective and efficient utilization of resources and energy. Upon entry into the stationary phase, transcription apparatus is modulated by replacing RNAP core-associated sigma (promoter recognition subunit) from growth-related RpoD to stationary-phase-specific RpoS. The anti-sigma factor Rsd participates for the efficient replacement of sigma, and the unused RpoD is stored silent as Rsd–RpoD complex. On the other hand, functional 70S ribosome is transformed into inactive 100S dimer by two regulators, ribosome modulation factor (RMF) and hibernation promoting factor (HPF). In this review article, we overview how we found these factors and what we know about the molecular mechanisms for silencing transcription apparatus and translation machinery by these factors. In addition, we provide our recent findings of promoter-specific transcription factor (PS-TF) screening of the transcription factors involved in regulation of the *rsd* and *rmf* genes. Results altogether indicate the coordinated regulation of Rsd and RMF for simultaneous hibernation of transcription apparatus and translation machinery.

Keywords: RNA polymerase sigma factor, anti-sigma factor (Rsd), ribosome, ribosome modulation factor, hibernation, stationary phase, *Escherichia coli* K-12

INTRODUCTION

Batch cultures under optimal laboratory conditions of the well-characterized model bacterium *Escherichia coli* in rich media at an optimum temperature (usually at 37°C, the temperature of host animals for enterobacterium *E. coli*) under sufficient supply of oxygen exhibit a progression of constant steady-state growth as measured by either counting of the viable cells or measuring the cell turbidity. Traditionally, the cell growth has been classified into three phases: non-replicative lag phase; replicative exponential phase; and stationary phase of replication cessation. The

growing-phase *E. coli* has long been used as a model organism relying on the belief that its laboratory culture is homogenous in cell populations. Most of our knowledge of modern molecular genetics such as the mechanisms and regulation of gene expression was established using such apparently homogenous planktonic cell cultures.

In contrast to the laboratory culture conditions, the conditions that allow steady-state bacterial growth are seldom found in nature. Instead, the lack of nutrients, accumulation of toxic waste compounds, and the influence of harsh environmental conditions such as lack of oxygen and pH change threaten the survival of *E. coli*. A variety of protection systems against such hazardous environments are induced for survival by changing the cell organization at both the molecular and cellular levels (Foster, 1999; Raivio, 2005; Battesti et al., 2011; Jin et al., 2012; Mehta et al., 2015). Under such a background, the focus in *E. coli* research is being shifted toward understanding the survival strategy of *E. coli* after growth cessation. Facing this research stage, *E. coli* is again recognized as a suitable model organism because of huge amounts of accumulated knowledge of *E. coli* such as the functions and regulation of the whole set of genes on its genome.

Upon entry into the stationary phase of laboratory *E. coli* cultures, a variety of morphological and physiological changes take place in individual cells. The growth phase-coupled changes in cell characteristics are associated with a change in expression pattern of the genome: most of the growth-related genes are turned off or leveled down, and, instead, a number of the genes needed for stationary-phase survival are expressed (for reviews, see Lowen and Hengge-Aronis, 1994; Ishihama, 1997; Ishihama, 1999). Overall level of genome expression decreased down to less than 10% of the level of exponential growth. The change in genome expression is mainly attributable to the changes in activity and specificity of gene expression system, including transcription apparatus and translation machinery in parallel with the structural reorganization of genome within the nucleoid (**Figure 1**). Upon entry into the stationary phase, unused excess cellular components are generally degraded for reuse as nutrients for survival. Both transcription apparatus and translational machinery are, however, stored without being degraded, and instead, their activity and specificity are markedly modulated for expression of the stationary-phase genes (referred to as "stationary genes" in this report). The major change of transcription apparatus is the replacement of the promoter-recognition subunit sigma from RpoD to RpoS through the aid of anti-sigma factor Rsd (regulator of sigma D) (Jishage and Ishihama, 1995) (**Figure 1**). On the other hand, 70S ribosome is converted into inactive 100S dimer with the aid of ribosome modulation factor (RMF) and hibernation promoting factor (HPF) (Maki et al., 2000; Ueta et al., 2005) (**Figure 1**). We found that these factors have been involved in detailed analyses of the regulatory roles of these factors (for reviews, see Wada, 1998; Ishihama, 1999; Ishihama, 2000; Yoshida and Wada, 2014). Here, we provide an overview of the molecular basis of genome expression system after the stationary phase, focusing on the simultaneous and coordinated hibernation of the transcription apparatus and the translation machinery.

Up to the present time, a set of anti-sigma factors have been identified, each sequestering each of all seven *E. coli* K-12 sigma factors (Hughes and Mathee, 1998; Helmann, 1999; Trevino-Quintanilla et al., 2013; Paget, 2015). Similar systems of the functional modulation of RNA polymerase (RNAP) are also known in bacteria other than *E. coli*, but the knowledge of regulatory functions of the whole set of sigma and anti-sigma factors is best known for *E. coli* (for details, see *Hibernation of the Transcription Apparatus*). Likewise, the factors for ribosome silencing differ between *E. coli* and other bacteria. For instance, non-gamma proteobacteria form 100S ribosome but lack RMF and contain long HPF homologues (Ueta et al., 2008; Yoshida and Wada, 2014) (for details see *Hibernation of the Translation Machinery*). As to the silencing of transcriptional apparatus and translational machinery, we focus on the well-characterized *E. coli* K-12 systems in this review.

GROWTH PHASE-COUPLED CHANGES IN CELL CHARACTERISTICS

Discontinuous Change of the Cell Buoyant Density

Upon entry into the stationary phase of laboratory *Escherichia coli* cultures, a variety of morphological and physiological changes take place in individual cells, including decrease in cell size, alteration in cell shape, compaction of nucleoid, changes in cell wall organization, and alterations in cytoplasm compositions (Roszak and Colwell, 1987; Kolter et al., 1993; Huisman et al., 1996). The synchronization of cell growth is disturbed, supposedly due to difference in microenvironment, and accordingly, the stationary-phase culture includes a mixture of heterogeneous cell populations including dead cells. The level and mode of cell heterogeneity differ depending on the culture conditions or factors affecting growth retardation the (Ferenci, 2001; Stewart and Franklin, 2008; Martinez-Antonio et al., 2012; Serra and Hengge, 2014; Pletnev et al., 2015). Upon entry into the stationary phase, the cell wall becomes thicker while the cytoplasm becomes condensed. In parallel, a variety of changes have been recognized for the cell characteristics, including the increase of unsaturated fatty acids in membrane, the increase of osmoprotective solutes such as trehalose and glycine betaine in cytoplasm, the accumulation of storage compounds such as glycogen and polyphosphate, and the decrease in polyamines (Roszak and Colwell, 1987; Kolter et al., 1993; Huisman et al., 1996; Ishihama, 2000). The nucleoid becomes more compact by replacing the DNA-binding proteins, for instance, from Fis in the log-phase to Dps in the stationary phase (Talukder et al., 1999; Ishihama, 2009). The DNA superhelicity, however, decreases in the stationary phase (Jaworski et al., 1991; Kusano et al., 1996).

For physical separation of heterogeneous cell populations, we succeeded in separating *E. coli* cell populations using centrifugation through gradients of polyvinylpyrrolidone-coated silica Percoll that protects the cells from toxic effects of silica (Makinoshima et al., 2002; Makinoshima et al., 2003). Due to the low viscosity of Percoll, materials as large as marker beads and bacterial cells quickly sediment to positions characteristic

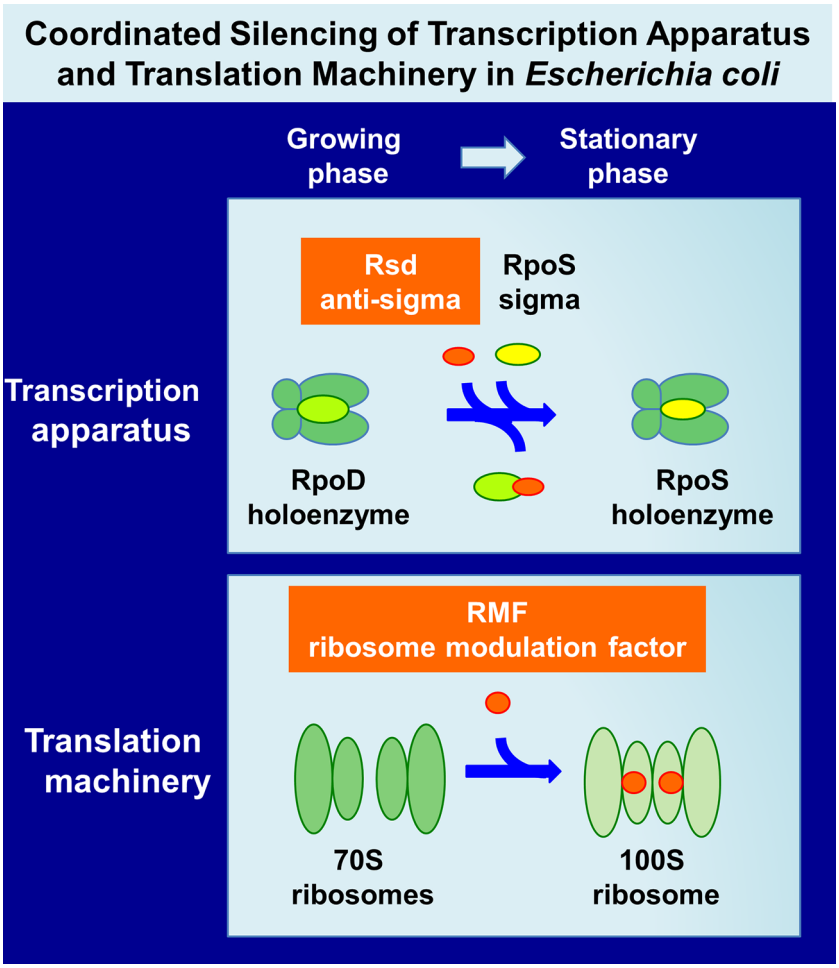

FIGURE 1 | Hibernation of transcription apparatus and translation machinery in *Escherichia coli* K-12. Upon entry of *E. coli* growth into the stationary phase, RNAP RpoD becomes silent through binding of anti-sigma factor Rsd onto the RpoD region-4 (promoter -35 recognition site) (Jishage and Ishihama, 1998; Jishage et al., 2001) while functional 70S ribosomes are converted to inactive 100S dimers through association with RMF (Wada et al., 1990; Wada, 1998) and HPF (Ueta et al., 2008; Yoshida and Wada, 2014). Here, we describe the coordinated regulation of two key regulators, Rsd and RMF, in *E. coli* K-12. The binding targets and binding sites of these two regulators on RNAP and ribosomes are described in text and also in **Figure 6**. Other factors involved in these processes are also described in text. RNAP, RNA polymerase; RMF, ribosome modulation factor; HPF, hibernation promoting factor.

of their densities. Exponential phase cultures of *E. coli* K-12 formed at least five discrete even though the density difference is within a narrow range (**Figure 2A**). This minor heterogeneity might correspond to the difference in the cycle of cell division (Kubitschek et al., 1983; Koch, 1996). In contrast, the stationary-phase cultures formed more than 10 bands, all exhibiting increased densities than the log-phase cultures (**Figure 2A**). A number of factors should influence the cell density, such as the cell volume, the chemical composition of cells, and the content of free water. One of the unexpected findings is the growth phase-coupled discontinuous transition of *E. coli* cell density. Even if the growth phase-coupled changes in molecular events are continuous, the overall cell characteristics change in discontinuous fashion as detected by the buoyant density. We concluded that the overall state of cell morphology and/or physiology of *E. coli* cells changes in discontinuous fashion during the growth transition from the log phase to the stationary phase.

A number of stationary genes have been identified by transcriptome and proteome analyses (Franchini et al., 2015; Sanchuki et al., 2017; Caglar et al., 2018). At present, however, we have only fragmentary knowledge on the expression order and the physiological roles of these stationary genes. We realized that the discontinuous change in cell buoyant density is a good marker for identification of the genes involved in each step of the cell differentiation during the transition of cell growth from exponential to stationary phase. We then subjected more than 200 single-gene-knockout mutants from the Keio collection (Baba et al., 2006; Yamamoto et al., 2009) to Percoll gradient centrifugation. Some mutants exhibited altered distribution (see **Supplemental Figure S2** for protein distribution), mostly defective in the density increase even after prolonged centrifugation. For instance, the density increase was found to be impaired at an early step for a mutant *E. coli* with the disrupted *rpoS* gene, which encodes RpoS sigma, the key player of stationary gene transcription (**Figure 2B**).

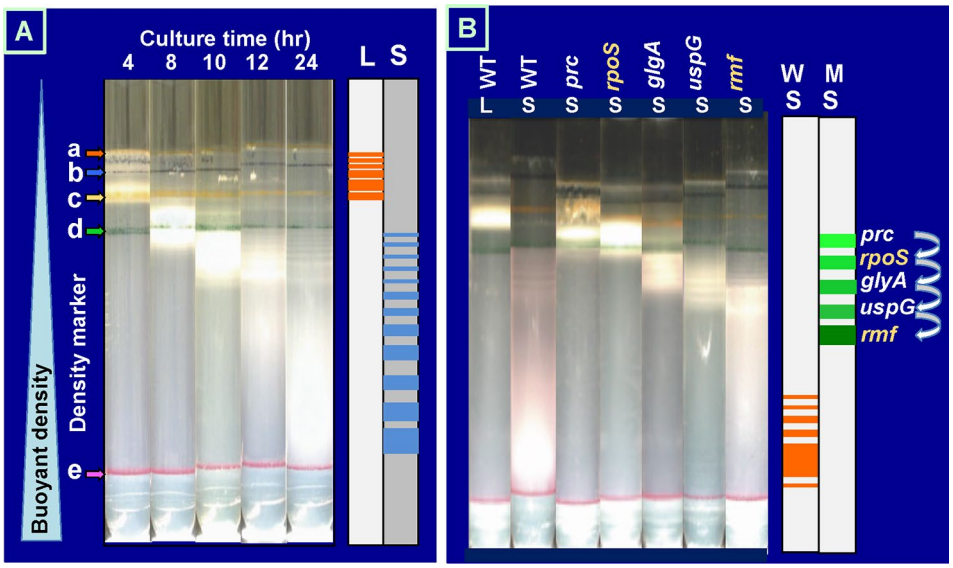

FIGURE 2 | Growth phase-dependent discontinuous increase of cell buoyant density of *Escherichia coli* K-12. **(A)** *E. coli* W3110 was grown in LB medium at 37°C with shaking. At various times, an aliquot of cell suspension was subjected to Percoll gradient centrifugation for 1 h at 20,000 rpm at 4°C in a Beckman SW40Ti rotor (Makinoshima et al., 2002; Makinoshima et al., 2003). The location of marker beads is indicated on the left: a, 1.035 g/ml; b, 1.074 g/ml; c, 1.087 g/ml; d, 1.102 g/ml; e, 1.119 g/ml. **(B)** *E. coli* wild-type BW25113 and its single-gene knockout mutants were grown in LB for 4 (L) or 24 h (S) and subjected to Percoll gradient centrifugation. The increase in cell buoyant density was interfered for these mutants, remaining at specific positions as indicated on the right. LB, lysogeny broth.

RpoS was found to be needed at the early stage of the cell density increase (for details, see next chapter). The interruption of density increase was observed for the genes not directly related to transcription. For instance, mutants defective in RpoF and RpoN exhibited essentially the same centrifugation pattern with that of wild-type *E. coli* K-12. In contrast, the density increase stopped for the mutant lacking the *rmf* gene at a step later than that for RpoS sigma, indicating that the ribosome dimerization takes place after expression of RpoS-dependent genes. Afterward, the density increase is interrupted for the mutant lacking universal stress protein (UspG) (**Figure 2B**). RMF is required for hibernation of ribosomes through conversion of functional 70S monomer to inactive 100S dimer (for details, see below) (Wada, 1998; Yoshida and Wada, 2014), while UspG is needed for cell–cell interaction in biofilm formation in the stationary phase (Nachin et al., 2005). The stop order of buoyant density increases for the *uspG* and *rmf* mutants agrees well with the order of maximum expression of UspG and RMF in wild-type *E. coli* (see **Figure 3**).

Growth-Dependent Change of the Protein Expression Pattern

As noted above, the pattern of genome expression in the stationary-phase changes for adaptation and survival as measured by genome-wide expression patterns of mRNA and protein products using the modern omics systems. In this section, we focus on the expression and degradation of the whole set of stationary proteins during the prolonged culture after the stationary phase up to 8 days. For protein separation and identification, we employed the radical-free highly reducing (RFHR) system of two-dimensional (2D) gel electrophoresis (for details, see Wada, 1986a; Wada, 1986b). The RFHR method allowed fine resolution of proteins on 2D gels, minimizing artificial spots generated through intra-molecular and inter-molecular Cys–Cys bridging under oxidation circumstances. The level of each protein on the RFHR 2D gel pattern can be determined by measuring the density of stained protein spot (**Supplemental Figure S1**). For the analysis of stationary proteins, we used *E. coli* K-12 AD202 strain lacking the *ompT* gene encoding outer membrane protease 7, which exhibits strong protein hydrolysis activity during cell lysate preparation once liberated from the outer membrane. In the experiments shown in **Figure 3**, cells were harvested at various times up to day 8. Under the culture conditions employed, the viability decreased gradually to less than 10% at day 8 (**Figure 3**, inset). The whole cell lysates were fractionated by centrifugation into CD (insoluble cell debris) and CE (cell extract supernatant fraction), which were then fractionated into CR (crude ribosome fraction) and PRS (post ribosomal supernatant fraction) (for details, see **Figure 3** legend). The nature of each protein spot on RFHR 2D gel could be determined after protein sequencing and/or mass spectroscopy. After repeating RFHR analysis thoroughly, a total of more than 650 protein spots were identified, of which a total of 65 appeared or markedly increased after the stationary phase. These proteins were detected in three cellular fractions: 31 in RPS, 30 in CD, and 4 in CR (**Supplemental Table S1**). Up to the present time, a total of 48 spots have been identified, but 17 remained unidentified.

The RFHR system is in particular useful for analysis of small proteins, allowing the identification of these small-sized ribosome-associated proteins. The CR (crude ribosomal) fraction contained the newly identified 50S proteins, L35 (RpmI)

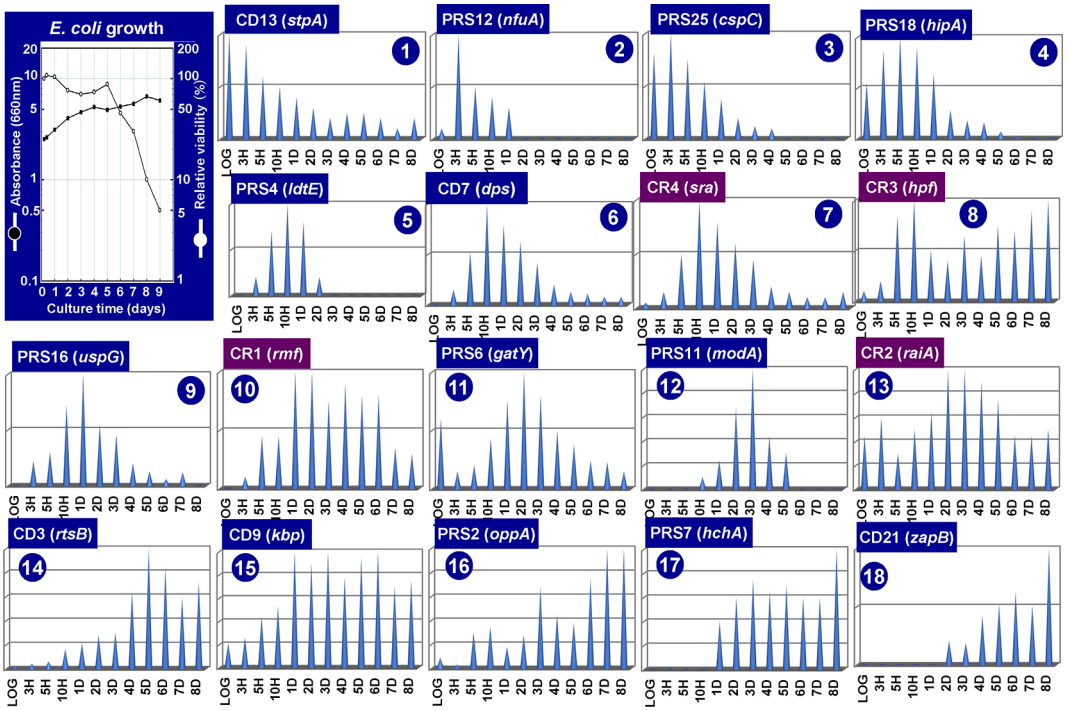

FIGURE 3 | Growth phase-dependent synthesis of 18 representative stationary proteins in *Escherichia coli* K-12. *E. coli* K-12 AD202 was grown in minimal medium E (Vogel and Bonner, 1956) containing 2% peptone at 37°C. The cell growth was monitored for 10 days by measuring the turbidity at 660 nm and by counting viable cells as shown in the inset. Aliquots of the culture were harvested at the indicated time (X-axis), and the cell lysates were fractionated into CD (insoluble cell debris), CE (cell extract supernatant), CR (crude ribosome), and PRS (post ribosomal supernatant) fractions. All these fractions prepared at each time point was subjected to RFHR 2D gel system, and the stained protein spots were measured by densitometry. The relative levels (Y-axis) are shown at each culture time (X-axis) for a total of 18 representative stationary proteins. The proteins shown under purple background indicate those involved in the hibernation of ribosomes.

and L36 (RpmJ) (Wada and Sako, 1987), and 30S protein S22 (Sra or RpsV) (Izutsu et al., 2001), leading to make the complete list of 54 r-proteins in *E. coli* K-12. Besides, some ribosome-related proteins were included in the CR fraction such as RMF, RaiA (renamed YfiA), and HPF (renamed YhbH), which all are involved in ribosome hibernation; for details, see *Hibernation of the Translation Machinery*.

The CD fraction recovered in the pellet fraction after low-speed centrifugation includes a total of 30 proteins tightly associated with cell wall and membrane. Stationary-phase-specific nucleoid proteins Dps and StpA were recovered in this CD fraction in agreement with the tight association of stationary-phase nucleoid with the cell membrane (Ishihama, 2009). Most of stress-response gene products in this CD fraction such as SlyD (chaperone with peptidyl-prolyl *cis-trans* isomerase activity) and StpA (H-NS-like nucleoid protein with RNA chaperone function), and two of six *E. coli* UspGs, UspD and UspG. All these proteins are involved in repair and refolding of RNAs and proteins (see **Supplemental Table S1**). The PRS fraction includes a total of 31 soluble stationary proteins, which all migrated in neutral to acidic regions on 2D (see **Supplemental Figure S1**). Most of these soluble proteins are involved in stationary-phase-specific metabolism, supposedly for redirection of metabolic circuits after prolonged culture in the absence of sufficient nutrients.

The level of stationary-phase proteins was measured throughout the culture up to day 8 (**Figure 3**), and the relative distribution is aligned in the order of appearance time throughout the 8-day culture (**Figure 4**). About half of the stationary-phase proteins appeared at specific time and soon disappeared, exhibiting a relatively narrow pattern of appearance in the stationary phase, but some other stationary proteins distributed in rather wide range of the stationary phase even though the distribution pattern between three subcellular fractions change. It should be noted that some stationary-phase proteins are detected in more than two fractions and exhibited culture time-dependent shift of distribution such as RPS-to-CD for GatY, RbsB, SlyD, UspD, ZapB, YdcH, and YibJ (see **Table 1**). The final deposition of these soluble proteins could be in the cell membrane and cell wall after prolonged culture. One exceptional distribution pattern was observed for RaiA, which showed a culture time-dependent alteration of distribution among all three fractions, CR, PRS, and CD (see **Table 1**), supposedly reflecting to its role in ribosome hibernation (see below).

Furthermore, it is interesting to note that even in the last day 8, expressions of some stationary-phase proteins are synthesized, including HchA (protein/nucleic acid deglycase), Mdh (malate dehydrogenase), GuaB (inosine 5′-monophosphate dehydrogenase), and ZapB (cell division factor). HchA is involved in repair of glyoxal- and methylglyoxal-glycated proteins (Mihoub

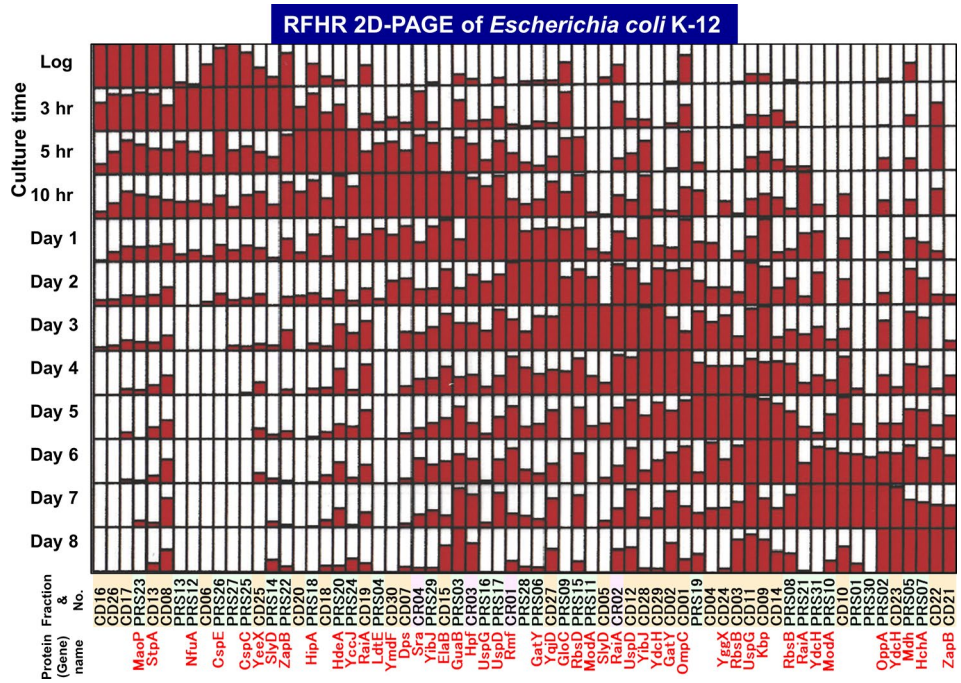

FIGURE 4 | Growth phase-dependent expression patterns of a total 65 stationary proteins in *Escherichia coli* K-12. The growth phase-dependent synthesis was measured for a total 65 stationary proteins. The relative level of synthesis from log phase (3-h culture) to day 8 is shown for all 65 proteins. The maximum level is shown by filling the day column with full red color. Spot numbers listed in **Table 1** are shown on the horizontal axis, with colors indicating fraction type (green: PRS, orange: CD, and magenta: CR). The protein products so far identified are shown in red below the corresponding spot numbers.

et al., 2015) and nucleic acids (Richarme et al., 2017). The *mdh* gene is also organized a network of genes, which facilitate stress-induced mutagenesis (Al Mamun et al., 2012). ZapB plays, together with ZapA, a role in organization and dynamics of the repaired genome in resting cells and independent of the Min system (Bailey et al., 2014; Mannik et al., 2016). Under stressful conditions unfavorable for *E. coli* growth, mutation rate increases for adaption and survival (Foster, 1999; Zinser and Kolter, 2004; Saint-Ruf et al., 2007). These 8-day proteins might be involved in repair of the genome and damaged proteins.

Both the sequential increase in cell buoyant density and the sequential synthesis of stationary-phase proteins are apparently under a single pathway, but it should be noted that the pathway for entry into the stationary phase is multiple. During the prolonged culture, the heterogeneity in the cell population should also be amplified due to generation of various types of cells on different pathways, such as persister cells, mutant cells, and dead cells (Roszak and Colwell, 1987; Kolter et al., 1993; Huisman et al., 1996; Ishihama, 1999).

GROWTH PHASE-COUPLED ALTERATIONS IN GENE EXPRESSION APPARATUS

Hibernation of the Transcription Apparatus

Upon entry into the stationary phase, the level of transcription decreases to less than 10% of that in the log phase (Ishihama, 2000). For this marked reduction in transcription pattern, the modulation of the promoter selectivity of RNAP is the major mechanism through the replacement of sigma subunit (the promoter recognition factor). In *Escherichia coli* K-12, seven different species of the sigma subunit exist, each recognizing a specific set of promoters (Ishihama, 1988; Ishihama, 2010). Transcription of the genes highly expressed in exponential growth phase is carried out by the RNAP holoenzyme containing RpoD, while RpoS is a key factor in the change in genome expression during growth transition from the exponential growth phase to the stationary phase (Lowen and Hengge-Aronis, 1994; Ishihama, 2010; Ishihama, 2012). We have measured the intracellular level of each sigma subunit at various phases of cell growth (**Figure 5A**). In exponentially growing cells of *E. coli* K-12, a significant level was detected only for three sigma factors, RpoD for growth-related genes, RpoN for nitrogen-assimilation genes, and RpoF for flagella-chemotaxis genes (Ishihama et al., 1976; Kawakami et al., 1979; Jishage and Ishihama, 1995). The concentration of RpoD is maintained at a constant level of 500–700 molecules per genome from log to stationary phase. The log-phase cells contain 1,500 to 2,000 molecules of RNAP core enzyme per genome, but about two-third are involved in transcription cycle (Ishihama and Fukuda, 1980; Ishihama, 2000). After transcription initiation, RpoD sigma is released, and the majority of free RNAP core might be associated with RpoD sigma, forming the RpoD holoenzyme.

TABLE 1 | Proteins Expressed During Prolonged Culture of *Escherichia coli* K-12.

PRS		CD		CR						
2D spot	Max stage	2D spot	Max stage	2D spot	Max stage	Gene		Map	pI/Size (aa)	Function
PRS16	Day 1	CD11	Log			*uspG*	*ybdQ,yzzU*	13.79	6.03/142	universal stress protein G
RPS26	Late-log (3 h)					*cspE*	*msmC*	14.16	8.09/69	transcription antiterminator/RNA stability regulator CspE
PRS10/11	Day-3 Day-7					*modA*		17.12	7.81/257	periplasmic molybdate transporter protein
		CD07	10 h			*dps*	*pexB,vtm*	18.27	5.70/167	stationary-phase nucleoid protein/Fe-binding storage protein
PRS09	Day 3					*gloC*	*ycbL*	21.19	4.95/215	hydroxyacylglutathione hydrolase;methylglyoxal degradation
				CR01	Day 1 and 2	*rmf*		21.87	10.86/55	ribosome modulation factor
PRS24	Late-log (5 h)					*yccJ*		22.97	4.70/75	PF13993 family protein YccJ
		CD30	10 h			*ymdF*		23.00	9.87/57	stress-induced acidphilic repreak motifs-containing protein
RPS02	Day 7 and 8					*oppA*		24.04	6.05/543	periplasmic oligopeptide transporter protein
PRS31	Day 2	CD23	Day 7 and 8			*ydcH*		32.29	9.30/74	uncharacterized protein
PRS18	Late-log (5 h)					*hipA*		34.28	8.26/440	serine/threonine kinase HipA; regulator with hipB
				CR04	10 h	*sra*	*rpsV*	35.52	11.04/45	30S ribosomal protein S22
PRS04	10 h					*ldtE*	*ynhG*	37.87	9.42/334	L,D-transpeptidase
PRS25	Late-log (3 h)					*cspC*	*msmB*	41.08	6.54/69	cold-shock stress protein CspC
PRS07	Day 8					*hchA*	*yedU,yzzC*	43.86	5.63/283	protein/nucleic acid deglycase; Hsp32 moleccular chaperone
		CD25	Late-log (3 h)			*yeeX*		44.79	9.30/109	DUF496 domain-containing protein
PRS06	Day 2	CD02	Day 4			*gatY*	*yegF*	46.91	5.87/284	tagarose-1,6-dibphosphate aldolase
		CD01	Day 4			*ompC*	*meoA,par*	49.82	4.58/367	outer membrane protein C pore for passive difusion
		CD15	10 h and Day 1			*elaB*	*yfbD*	51.34	5.35/101	tail-anchored inner membrane protein
PRS03	Day 7					*guaB*		56.60	6.02/486	Inosine 5'-monophosphate dehydrogenase; GMP symthesis GMP symthesis
PRS21	Day 7	CD19	10 h	CR02	Day 2 and 3	*raiA*	*yfiA*	58.88	6.19/113	stationary-phase translation inhibitor/ribosome stability factor
		CD13	Log			*stpA*	*hnsB,rsv*	60.19	7.95/134	nucleoid protein StpA with RNA chaperone activiry
		CD09	Day 6			*kbp*	*ygaU,yzzM*	60.24	5.67/149	K+ binding protein
		CD24	Day 5			*yggX*		66.78	5.91/91	Fe2+-tracking protein; oxidative damage protect Fe-S protein
		CD27	Day 2			*yqjD*		69.91	9.06/101	ribosome- and membrane-associated DUF-domain protein
				CR03	10 h	*hpf*	*yhbH*	72.01	6.50/95	ribosome hibernation-promoting factor; RpoN modulation protein
RPS05	Day 8					*mdh*		72.81	5.61/312	malate dehydrogenase
PRS22	Late-log (3 and 5 h)	CD21	Day 8			*zapB*	*yiiU*	75.71	4.69/81	cell division factor ZapB
RPS17	Day 1	CD12	Day 3 and 7			*uspD*	*yiiT*	75.82	6.37/142	universal stress protein D
RPS08	Day 6	CD03	Day 5			*rbsB*	*priB,rbsP*	79.62	6.85/296	periplasmic ribose transperter protein
RPS15	Late-log (3 h)					*rbsD*	*rbsP*	79.70	5.93/139	D-ribose pyranase; sugar-binding protein
RPS29	10 h	CD28	Day 3			*yibJ*		83.35	5.00/?	RHA domain-containing protein YibJ
RPS23	Log					*maoP*	*yifE*	85.06	6.09/112	macrodomain Ori protein
RPS20	Late-log (5 and 6 h)					*hdeA*	*yhhC,yhiB*	85.74	5.06/110	periplasmic acid stress chaperone HdeA
PRS12	Late-log (3 h)					*nfuA*	*gntY,yhbI*	88.12	4.52/191	iron-sulfur cluster carrier protein; gluconate transporter
PRS14	Late-log (3 h)	CD05	Day 3			*slyD*		89.57	4.86/196	FKBP-type pepridyl-prolyl cis-trans isomerase
PRS01	Day 7					*X*				
PRS13	Late-log (3 h)					*X*				

(Continued)

TABLE 1 | Continued

PRS		CD		CR					
2D spot	Max stage	2D spot	Max stage	2D spot	Max stage	Gene	Map	pI/Size (aa)	Function
PRS19	Day 5					X			
PRS27	Late-log (2 h)					X			
PRS28	Day 2					X			
RPS30	Day 7					X			
		CD04	Day 5			X			
		CD06	Late-log (3 h)			X			
		CD08	Log			X			
		CD10	Day 5 and 7			X			
		CD14	Day 6			X			
		CD16	Log			X			
		CD17	Log			X			
		CD18	Late-log (5 h)			X			
		CD20	Late-log (5 hr)			X			
		CD22	Day 8			X			
		CD26	Log			X			

(green: PRS, orange: CD, and magenta: CR; see legends of **Figure 3** *and* **4***).*

RpoS sigma is needed for transcription of stationary-phase genes. The level of RpoS starts to increase after the mid-log phase and reaches to the maximum level of about the half the level of RpoD in the stationary phase (**Figure 5B**) (Jishage and Ishihama, 1995; Jishage et al.,1996). The level of core enzyme is under the autogenous control, thereby keeping the constant level of about 2,000 molecules per genome throughout cell growth (Ishihama, 2000). In contrast, the combined level of all seven sigma factors is about two folds the level of the core enzyme, and we then proposed the "sigma competition" model (Jishage and Ishihama, 1998; Maeda et al., 2000). Since the level of RpoD was always higher than RpoS even after prolonged culture, we doubted whether RpoD is still functional in the stationary phase. As an attempt to examine this possibility, we analyzed proteins associated with RpoD at various phases of cell growth and discovered the association of a novel protein Rsd (regulator of sigma D) (Jishage and Ishihama, 1998; Jishage and Ishihama, 1999), which forms a complex with RpoD for interfering with its sigma function. The level of Rsd starts to increase upon entry into the stationary phase, finally reaching to the level of 60 to 80% of RpoD (**Figure 5B**), implying that most of RpoD stays non-functional in the stationary phase through formation of RpoD-Rsd complex. As a result, the core enzyme becomes available for association of the stationary-specific RpoS sigma (Jishage and Ishihama, 1998; Mitchell et al., 2007). The anti-sigma factor Rsd binds to the RpoD domain-4 that is involved in recognition of the promoter -35 signal (Dove and Hochschild, 2001; Jishage et al., 2001; Mitchell et al., 2007) (**Figure 5C**). Crystal structure of Rsd–RpoD complex supports this conclusion (Patikoglou et al., 2007). The affinity of Rsd to free RpoD is high, and in the presence of high concentrations of Rsd, it also binds to the core-associated RpoD (Ilag et al., 2004; Westblade et al., 2004). After sequestering RpoD into Rsd–RpoD complex, the free core enzyme could be used for formation of RpoS holoenzyme, thereby allowing transcription of stationary genes.

Based on these findings, we proposed the "sigma competition" model, in which the anti-sigma factor plays a regulator in replacement of RNAP-associated sigma for an efficient switching of its promoter selectivity (Jishage and Ishihama, 1999; Maeda et al., 2000; Mitchell et al., 2007). Along this line, it should be noted that the anti-sigma factors have been identified for all seven sigma factors of *E. coli* K-12 and widely in other bacteria (Hughes and Mathee, 1998; Helmann, 1999; Trevino-Quintanilla et al., 2013; Paget, 2015). To confirm the "sigma competition" model for control of the promoter selectivity of RNAP, we further compared the binding affinity *in vitro* of all seven sigma factors to the same core enzyme (Maeda et al., 2000). In the presence of a fixed amount of RpoD, the level of RpoD holoenzyme formation increased linearly with the increase in core enzyme level. Mixed reconstitution experiments in the presence of a fixed amount of the core enzyme and increasing concentrations of an equimolar mixture of all seven sigma subunits indicated that the core binding is the strongest for RpoD sigma, followed by RpoN, RpoF, RpoE, FecI, and RpoS in decreasing order. The order of core binding activity was also confirmed by measuring the replacement of one core-associated sigma by another sigma subunit. Since the intracellular level of core enzyme is virtually constant, the model of sigma replacement relies solely on changes in the intracellular concentrations of seven sigma subunits (Ishihama 2000; Ishihama 2010).

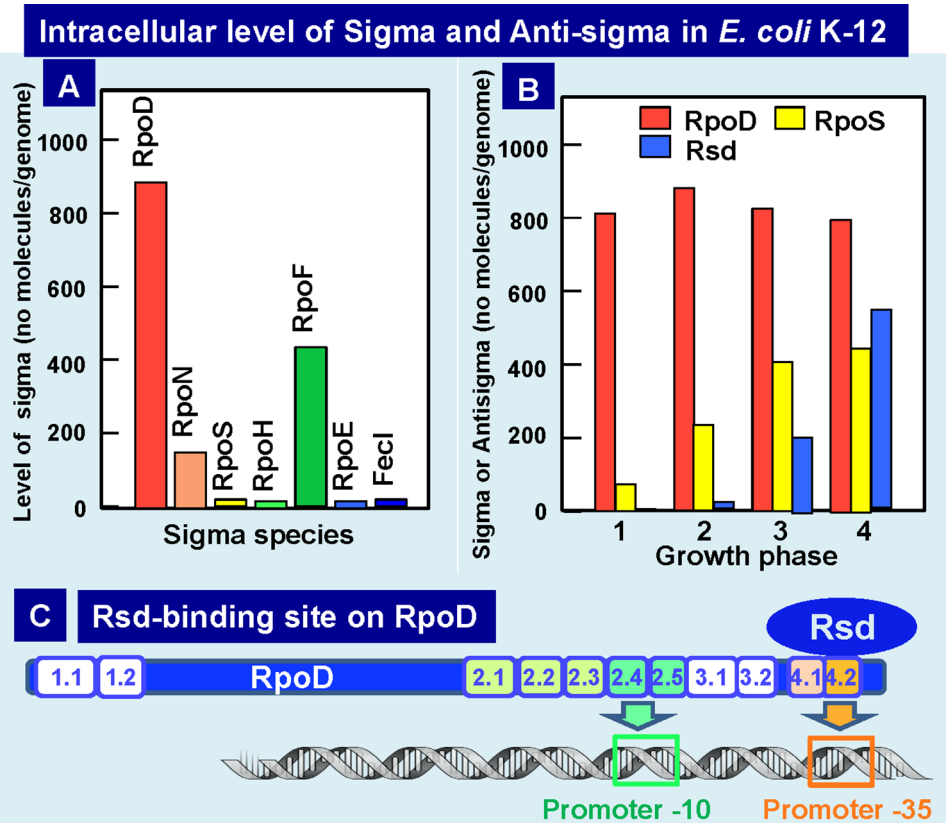

FIGURE 5 | Intracellular levels of sigma factors and anti-RpoD sigma (Rsd). **(A)** Intracellular levels of all seven sigma factors in exponential phase *E. coli* K-12 was determined by Western blot analysis with use of specific antibodies (Jishage and Ishihama, 1995; Jishage et al., 1996). **(B)** Intracellular levels of growth-related RpoD sigma, stationary-phase-specific RpoS sigma, and anti-RpoD sigma Rsd were determined at various growth phases of *E. coli* K-12 (Jishage and Ishihama, 1998; Jishage and Ishihama, 1999). **(C)** The contact site of anti-sigma factor Rsd on the growth-related RpoD sigma was determined to be located within RpoD region-4 (promoter -35 recognition site) by using the contact-dependent cleavage sites by Rsd-tethered iron-*p*-bromoacetamidobenzyl EDTA by analysis of the complex formation between Ala-substituted σ^{70} and Rsd (Jishage and Ishihama, 2001). Rsd-binding to RpoD region-3 leads to silencing RpoD function.

Besides RpoD sigma, Rsd was found to interact with HPr, a phosphocarrier component of PEP-dependent sugar-transporting phosphotransferase system (PTS), thereby interfering with anti-sigma activity (Park et al., 2013). Recently Rsd was also found to interact with SpoT and stimulates its hydrolysis activity of magic spot (p)ppGpp (Lee et al., 2018). The SpoT activity is, however, antagonized by dephosphorylated HPr, which generally interacts with a large number proteins and regulate wide varieties of carbon and energy metabolism (Rodionova et al., 2017). These observations altogether indicate the presence of a protein–protein interacting network between Rsd, HPr, and SpoT for interconnection between transcription and metabolism during the stationary phase.

Here, we propose the hibernation of growth-phase RNAP holoenzyme through conversion of RpoD sigma by Rsd anti-sigma factor. The RNAP core enzyme can then be used for assembly of RpoS holoenzyme for transcription of stationary-phase genes. It should be noted that excess free core enzyme, if present, should form transcriptionally inactive dimers or oligomers (Ishihama, 1990; Harris et al., 1995) for storage as in the case of yeast RNAP I (Fernandez-Tornero, 2018). The conversion of RpoD into the inactive RpoD-Rsd complex and the self-assembly of free core enzyme together contribute for silencing of the transcription apparatus during the stationary phase.

Hibernation of the Translation Machinery

Bacterial ribosomes are universally conserved ribonucleoprotein complexes, generally consisting of two asymmetric subparticles. In *E. coli* K-21, large (50S) and small (30S) subparticles associate with each other to form the functional 70S ribosomes. The 50S subparticle is composed of two species of rRNA (23S and 5S) and a total of 33 species of the ribosomal protein, referred to r-protein (L1 to L36), whereas the 30S subparticle is composed of 16S rRNA and a total of 21 species of r-proteins (S1 to S21) (Wada and Sako, 1987; Izutsu et al., 2001; Kaczanowska and Ryden-Aulin, 2007; Shajani et al., 2011). Under optimal laboratory culture conditions, *E. coli* grows exponentially with heavy consumption of energy and resources.

During this exponential phase, the ribosome profile detected by sucrose density gradient centrifugation (SDGC) includes 70S ribosomes as the major component and in addition, small amounts of 30S and 50S subparticles, and polysomes (**Supplemental Figure S3A**). These ribosomes are involved in the canonical ribosome cycle (initiation, elongation, termination, and recycling) of protein

synthesis (**Figure 6A**). Protein synthesis is the most energy demanding cellular process. The majority of metabolic energy is used for the formation of ribosomes (Maaloe and Kjeldgaard, 1966). Upon entry into the stationary phase, overall level of transcription decreases to less than 10% the level of log phase, yielding the superfluous translation machinery. The unused excess ribosomes are then converted into non-functional 100S ribosome dimers, the inactive stored form of ribosomes (**Supplemental Figure S3B** and **Figure 6B**) (Wada et al., 1990; Yoshida and Wada, 2014). The ribosome profile measured by SDGC includes a peak of 100S ribosomes besides the peak of 30S, 50S, and 70S ribosome (**Supplemental Figure S3B** and **Figure 6B**). The 100S ribosome is a dimer of 70S ribosomes, and inactive in translation (Wada et al., 1990; Wada et al., 1995). We then designated this stage of ribosome cycle, in which the ribosomes stay in inactive forms, for "Hibernation" (Yoshida et al., 2002).

The 100S ribosome of *E. coli* is formed by the binding of two factors, the RMF (Wada et al., 1990) and the HPF (Ueta et al., 2013). RMF alone leads only to the formation of 90S particle, which is an immature form of the 100S ribosome, suggesting that HPF is needed to convert this premature 90S particle to mature 100S ribosome (Ueta et al., 2005; Ueta et al., 2008; Ueta et al., 2013). The third protein associated with the stationary-phase ribosomes is RaiA (renamed YfiA), which interferes with the 100S dimer formation through competition with HPF binding (Maki et al., 2000; Ueta et al., 2005). Thus, two factors, HPF and RaiA, share the same binding site on the 100S ribosome and thus compete each other, thereby controlling the formation of 100S ribosomes. The binding sites of RMF and HPF investigated by several methods indicate the conformational changes of 30S subunits, thereby controlling the ribosome dimerization indirectly (Yoshida et al., 2002; Ueta et al., 2005; Yoshida and Wada, 2014; Beckert et al., 2018) (see **Figure 6**, right panel). Inactivation of the *rmf* gene leads to loss of viability in the stationary phase (Yamagishi et al., 1993), under acidic conditions (El-Sharoud and Niven, 2007) and upon exposure to heat shock (Niven, 2004). When the stationary-phase *E. coli* was transferred to nutrient-rich media, the disassembly of 100S ribosomes is rapid within 1 min (Aiso et al., 2005) for restart of protein synthesis (Yoshida and Wada, 2014). The

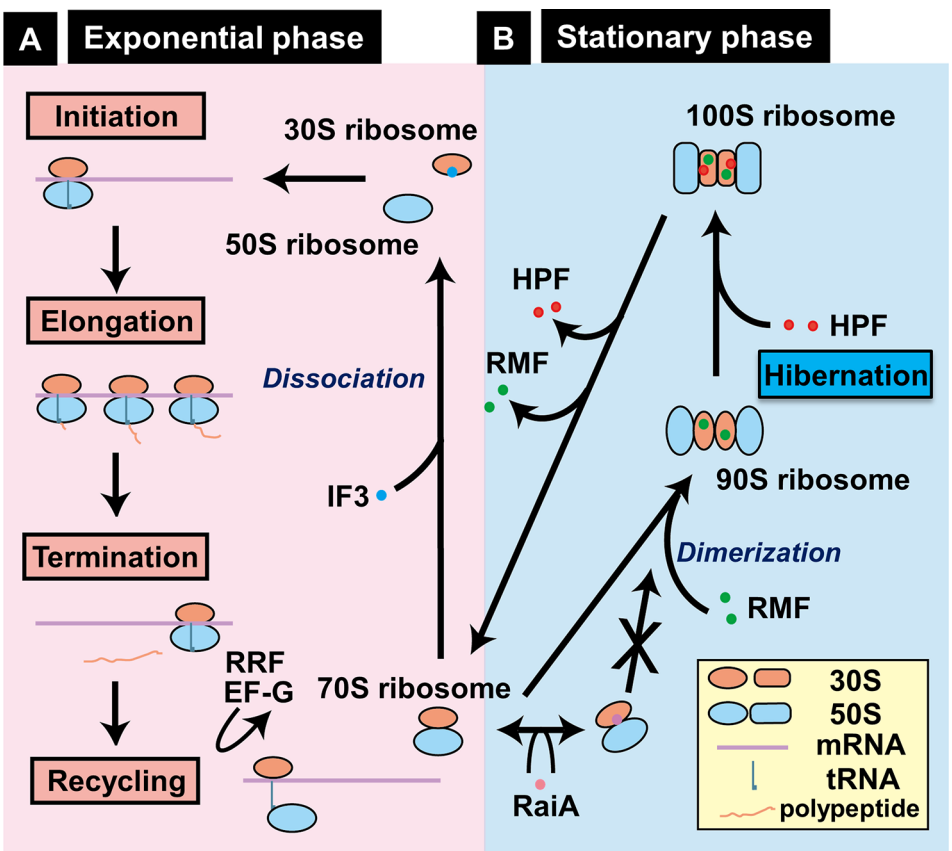

FIGURE 6 | Growth phase-coupled alteration of ribosomes in *Escherichia coli* K-12. **(A)** In exponentially growing bacterial cells, most ribosomes are involved in the functional cycle of protein synthesis, consisting of initiation, elongation, termination, and recycling. For initiation, 30S and 50S ribosomes bind to mRNA, forming functional 70S ribosomes on mRNA and ultimately leading to form polysomes. After termination, 70S ribosomes are dissociated into 30S and 50S subparticles for reutilization. **(B)** Upon entry into stationary phase, unused ribosomes are converted into functionally inactive 100S dimeric ribosomes by sequential binding of RMF and HPF in *E. coli* K-12, one of Gram-negative bacteria (Wada, 1998; Yoshida and Wada, 2014). We designated this process as "hibernation." Formation of 100S dimers is interfered by RaiA (renamed YfiA) (Ueta et al., 2005). The location of RMF on 30S ribosome is based on the recent cryo-electron micrography structure of 100S ribosome dimer (Beckert et al., 2018). By biochemical analyses, however, RMF was also indicated to bind 23S rRNA (Yoshida et al., 2004) and the peptidyl transferase center (Yoshida et al., 2002).

mechanism how RMF and HPF are removed from 100S ribosomes remains to be solved.

The ribosome hibernation is widespread but the factors involved in this process are different between bacteria (Ueta et al., 2008; Yoshida and Wada, 2014; Prossliner et al., 2018). *E. coli* and some γ-proteobacteria carry both the *rmf* and *hpf* genes, but many other bacteria have only the *hpf* gene or its homologue devoid of the *rmf* gene (Ueta et al., 2008). In bacteria carrying a long-type HPF homologue, the ribosome dimerization takes place in the absence of RMF (Ueta et al., 2013; Akanuma et al., 2016). *E. coli* forms 100S ribosomes only in their stationary growth phase, but in Gram-positive bacteria such as *Bacillus subtilis*, 100S ribosomal dimers are formed throughout entire growth phases (Ueta et al., 2013; Puri et al., 2014; Akanuma et al., 2016), implying that the factors or conditions for ribosome dimerization are different between bacterial species.

In the case of the bacterial group having long HPF, several structures have been proposed for the ribosome dimer (For instance, Matzov et al., 2019). Accordingly, the 70S–70S interface within ribosome dimers appeared different from that of *E. coli* (Kato et al., 2010; Beckert et al., 2018). Nevertheless, N-terminal domain of long HPF is predicted to bind to the site overlapping with the tRNA-binding site as in the case of HPF in *E. coli*, suggesting that common mechanism of translational silencing exists between bacteria carrying long and short HPFs.

COORDINATED HIBERNATION OF TRANSCRIPTION APPARATUS AND TRANSLATION MACHINERY

The formation of transcriptional apparatus and translational machinery are tightly coupled and coordinated, showing the growth rate-dependent synthesis of RNAP core enzyme (Ishihama and Fukuda, 1980; Ishihama, 1988) and ribosomes (Nomura et al., 1984; Zengel and Lindahl, 1994), thereby keeping the ratio of 5~10 ribosomes per RNAP core to match effective translation of mRNA through formation of polysomes. For this purpose, multiple layers of regulation are involved such as the organization of genes for RNAP subunits and ribosomal proteins into single and same operons, and the autogenous regulation of synthesis of RNAP subunits and ribosomal proteins by excess and unused products. We then examined the possible coordination in the hibernation process between transcription apparatus and translation machinery. During

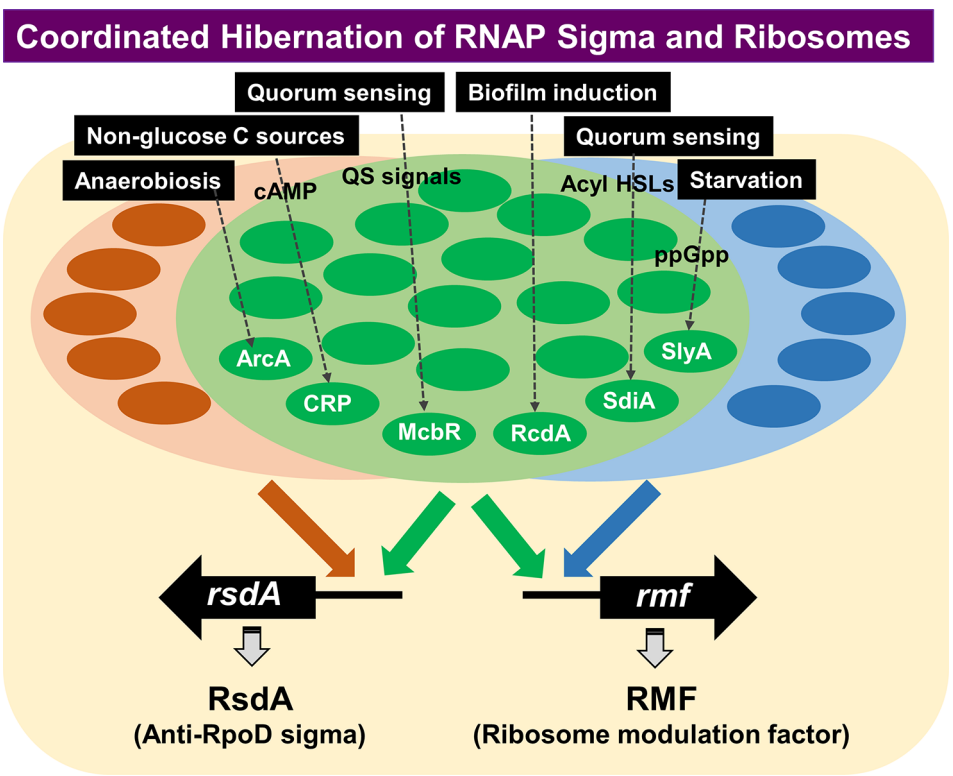

FIGURE 7 | PS-TF screening was performed for search of TFs involved in regulation of the *rsd* and *rmf* genes. A total of 74 TF species were found to bind to both the *rsd* and *rmf* promoter probes, although the binding affinity appeared different between these TFs (Yoshida et al., 2018). Besides these 74 TFs, some other TFs have been identified to bind only the *rsd* gene or the *rmf* gene, indicating independent regulation of the two genes under as yet unidentified conditions. Detailed analysis of the regulatory roles *in vitro* and *in vivo* was performed for the five representative stress-response TFs (ArcA, McbR, RcdA, SdiA, and SlyA) (Yoshida et al., 2018). ArcA was indicated to repress transcription of both *rsd* and *rmf* genes, while other four were suggested to activate both genes. gSELEX indicated that all these TFs regulate not only the *rsd* and *rmf* genes but also regulate a number of genes supposedly required for survival under stressful conditions. PS-TF, promoter-specific transcription factor.

the growth transition of *Escherichia coli* from log to stationary phase, the level of genome expression is reduced less than 10% the log-phase level and the pattern of genome expression (the species of expressed genes) is also markedly modulated. For this alteration, the transcription apparatus is altered by binding of anti-sigma factor Rsd to the RpoD sigma for sigma replacement with stationary-phase-specific RpoS (see above) while the translation machinery is modulated by binding of RMF and HPF to 70S ribosome to form the inactive 100S ribosome dimer (see above). Until recently, however, little was known how the expression of factors involved in hibernation of transcription apparatus and translation machinery is regulated. We have then performed a systematic search for TFs involved in regulation of the promoters of two key regulators, Rsd for hibernation of RNAP and RMF for hibernation of ribosomes, by using the newly developed promoter-specific transcription factor (PS-TF) screening system (Shimada et al., 2013; Yoshida et al., 2018).

Using *rsd* and *rmf* promoter probes and a total of about 200 purified TFs from *E. coli* K-12 W3110, we performed PS-TF screening (Yoshida et al., 2018). A total of 74 TF species (55 group A TFs and 19 group B TFs) were found to bind to both the *rsd* and *rmf* probes, although the binding affinity was different between these TFs (Yoshida et al., 2018), suggesting that both the *rmf* and *rsd* genes are under the control of multi-factor promoters (Ishihama et al., 2016). After repetition of PS-TF, we succeeded to focus on a total of 19 TFs, of which 9 (ArcA, CRP, CueR, McbR, NhaR, RcdA, SdiA, SlyA, and ZntR) have been experimentally confirmed to be involved in regulation *in vitro* and *in vivo* of both the *rsd* and *rmf* genes (Yoshida et al., 2018) (**Figure 7**). The synthesis of RMF is also under the control of ppGpp (Izutsu et al., 2001). Results altogether indicated the involvement of a common set of TFs, each sensing a specific but different environmental condition, in coordinated hibernation of the transcriptional apparatus and translational machinery for adaptation and survival under stressful conditions. Translation of RMF is stimulated by polyamines (Terui et al., 2010), which accumulates in the stationary phase (Igarashi and Kashiwagi, 2018).

Besides the large set of TFs with binding activity to both *rsd* and *rmf* probes, a small number of TFs bound only to either the *rsd* or *rmf* probe (**Figure 7**). This finding indicates the two key players for hibernation of transcription apparatus and translational machinery are regulated independently under certain specific conditions. These *rsd*- or *rmf*-specific TFs might be involved in independent regulation of either transcriptional apparatus or translational machinery under as yet unidentified specific environmental conditions. This review proves the initial stage of molecular basis of the hibernation of *E. coli*, focusing on the transcription apparatus and the translation machinery. The whole set of TFs involved in the regulation of *rsd* and *rmf* genes will be described elsewhere.

AUTHOR CONTRIBUTIONS

HY: data collection, writing, and figure preparation in Hibernation of the Translation Machinery and Coordinated Hibernation of Transcription Apparatus and Translation Machinery sections. TS: data collection and figure preparation in Discontinuous Change of the Cell Buoyant Density and Coordinated Hibernation of Transcription Apparatus and Translation Machinery. AW and YM: data collection, writing, and figure preparation in Growth-Dependent Change of the Protein Expression Pattern section. AI: design of this review article, data collection, writing, and figure preparation in Introduction, Discontinuous Change of the Cell Buoyant Density, Hibernation of the Transcription Apparatus, and Coordinated Hibernation of Transcription Apparatus and Translation Machinery. All authors read and approved the manuscript.

FUNDING

This study was supported by MEXT Cooperative Research Program of Network Joint Research Center for Materials and Devices to AI, and the MEXT-Supported Program for the Strategic Research Foundation at Private Universities to AI.

ACKNOWLEDGMENTS

We thank all those were involved in the works described in this article, in particular, H. Ogasawara (Shinshu Univ., Ueda), K. Yamamoto (Hosei Univ., Tokyo), M. Jishage (Rockefeller Univ. New York), N. Fujita (Tokyo Univ. of Agriculture, Tokyo), M. Ueta, and C. Wada (Yoshida Biol Lab, Kyoto). The *E. coli* strains used in this study were obtained from the Genetic Stock Center, National Institute of Genetics, Mishima.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.01153/full#supplementary-material

SUPPLEMENTAL FIGURE S1 | RFHR 2D gel patterns of *E. coli* proteins. *E. coli* K-12 AD202 was grown at 37 °C in minimal medium E containing 2% peptone. Cell lysates were prepared as described in text, and fractionated by centrifugation into CD (Insoluble cell debris), CE (supernatant cell extract), CR (crude ribosome), and PRS (post ribosomal supernatant) fractions. Soluble proteins of the CR [A], PRS [B], basic CD [C] and acidic CD [C] were subjected to RFHR (radical free and highly reducing) method of 2D gel analysis (Wada 1986a; Wada, 1986b). Gels were stained with CBB.

SUPPLEMENTAL FIGURE S2 | Contents of some growth-related proteins in cells fractionated by Percoll centrifugation. *E. coli* K-12 was grown at 37°C in LB medium for 12 hrs. Cell suspension was directly subjected to Percoll gradient centrifugation under the standard procedure (Makinoshima et al., 2002; Makinoshima et al., 2003). The gradient was fractionated and the content of some representative growth-related proteins in each fraction was determined by Western blot analysis with use of specific antibodies (Ishihama et al., 2014).

SUPPLEMENTAL FIGURE S3 | Ribosomes were subjected to sucrose density gradient centrifugation (SDGC) (Wada et al., 1990). **(A)** The pattern of ribosomes from exponential phase includes 70S ribosome as the major component and in addition, small amounts of 30S and 50S subparticles, and polysomes. **(B)** Ribosomes from stationary phase showed 100S dimeric ribosomes in addition to 30S, 50S and 70S ribosomes.

SUPPLEMENTAL TABLE S1 | Proteins Expressed During Prolonged Culture of Escherichia coli K-12 PRS (Post Ribosomal Supernatant) fraction

REFERENCES

Aiso, T., Yoshida, H., Wada, A., and Ohki, R. (2005). Modulation of mRNA stability participates in stationary-phase specific expression of ribosome modulation factor. *J. Bacteriol.* 187, 1951–1958. doi: 10.1128/JB.187.6.1951-1958.2005

Akanuma, G., Kazo, Y., Tagami, K., Hiraoka, H., Yano, K., Suzuki, S., et al. (2016). Ribosome dimerization is essential for the efficient regrowth of Bacillus subtilis. *Microbiology* 162, 448–458. doi: 10.1099/mic.0.000234

Al Mamun, A. A. M., Lombardo, M.-J., Shee, C., Lisewski, A. M., Gonzalez, C., Lin, D., et al. (2012). Identity and function of a large gene network underlying mutagenic repair of DNA breaks. *Science* 338, 1344–1348. doi: 10.1126/science.1226683

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., et al. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* 2, 2006.0008. doi: 10.1038/msb4100050

Bailey, M. W., Bisicchia, P., Warren, B. T., Sherratt, D. J., and Mannik, J. (2014). Evidence for divisome localization mechanisms independent of the Min system and SlmA in *Escherichia coli*. *PloS Genet.* 10, e1004504. doi: 10.1371/journal.pgen.1004504

Battesti, A., Majdaiani, N., and Gottesman, S. (2011). The RpoS-mediated general stress response in *Escherichia coli*. *Annu. Rev. Microbiol.* 65, 189–213. doi: 10.1146/annurev-micro-090110-102946

Beckert, B., Turk, M., Czech, A., Berninghausen, O., Beckmann, R., Ignatova, Z., et al. (2018). Structure of a hibernating 100S ribosome reveals an inactive conformation of the ribosomal protein S1. *Nat. Microbiol.* 3, 1115–1121. doi: 10.1038/s41564-018-0237-0

Caglar, M. U., Hockenberry, A. J., and Wilke, C. O. (2018). Predicting bacterial growth conditions from mRNA and protein abundances. *PloS One* 13, e0206634. doi: 10.1371/journal.pone.0206634

Dove, S. L., and Hochschild, A. (2001). Bacterial two-hybrid analysis of interactions between region 4 of the σ70 subunit of RNA polymerase and the transcription regulators Rsd from *Escherichia coli* and AlgQ from *Pseudomonas aeruginosa*. *J. Bacteriol.* 183, 6413–6421. doi: 10.1128/JB.183.21.6413-6421.2001

Dwek, R. D., Kobrin, L. H., Grossman, N., and Ron, E. Z. (1980). Synchronization of cell division in microorganisms by Percoll gradients. *J. Bacteriol.* 144, 17–21.

El-Sharoud, W. M., and Niven, G. W. (2007). The influence of ribosome modulation factor on the survival of stationary-phase *Escherichia coli* during acid stress. *Microbiology* 153, 247–253. doi: 10.1099/mic.0.2006/001552-0

Ferenci, T. (2001). Hungry bacteria – definition and properties of a nutritional state. *Environ. Microbiol.* 3, 605–611. doi: 10.1046/j.1462-2920.00238.x

Fernandez-Tornero, C. (2018). RNA polymerase I activation and hibernation: unique mechanisms for unique genes. *Transcription* 9, 248–254. doi: 10.1080/21541264.2017.1416267

Foster, P. L. (1999). Mechanisms of stationary phase mutation: a decade of adaptive mutation. *Annu. Rev. Genet.* 33, 57–88. doi: 10.1146/annurev.genet.33.1.57

Franchini, A. G., Ihssen, J., and Egli, T. (2015). Effect of global regulators RpoS and cyclic-AMP/CRP on the catabolome and transcriptome of *Escherichia coli* K12 during carbon- and energy-limited growth. *PloS One* 10, e0133793. doi: 10.1371/journal.pone.0133793 eCollection 2015.

Harris, S. J., Williams, R. C. Jr., and Lee, J. C. (1995). Self-association of *Escherichia coli* DNA-dependent RNA polymerase core enzyme. *Biochemistry* 34, 8752–8762. doi: 10.1021/bi00027a026

Helmann, J. D. (1999). Anti-sigma factors. *Curr. Opin. Microbiol.* 2, 135–141. doi: 10.1016/S1369-5274(99)80024-1

Hofmann, N., Wurm, R., and Wagner, R. (2011). The *E. coli* anti-sigma factor Rsd: Studies on the specificity and regulation of its expression. *PloS One* 6, e19235. doi: 10.1371/journal.pone.0019235

Hughes, K. T., and Mathee, K. (1998). The anti-sigma factors. *Annu. Rev. Microbiol.* 52, 231–286. doi: 10.1146/annurev.micro.52.1.231

Huisman, G. W., Siegele, D. A., Zambrano, M. M., and Kolter, R. (1996). Morphological and physiological changes during stationary phase, in *Escherichia coli and Salmonella*. Ed. Neidhardt, F. C. (Washington, DC: American Society for Microbiology Press), 1672–1682.

Igarashi, K., and Kashiwagi, K. (2018). Effect of polyamine on protein synthesis and growth. *J. Biol. Chem.* 293, 18702–18707. doi: 10.1074/jbc.TM118.003465

Ilag, L. L., Westblade, L. F., Deshayes, C., Kolb, A., Busby, S. J. W., and Robinson, C. V. (2004). Mass spectrometry of *Escherichia coli* RNA polymerase: Interactions of the core enzyme wtih σ^{70} and Rsd protein. *Structure* 12, 269–275. doi: 10.1016/j.str.2004.01.007

Ishihama, A., Taketo, M., Saitoh, T., and Fukuda, R. (1976). Control of formation of RNA polymerase in *Escherichia coli*, in RNA Polymerase. Ed. Chamberlin, M., and Losick, R. (New York, Cold Spring Harbor Laboratory Press), 485–502.

Ishihama, A., and Fukuda, R. (1980). Autogenous and post-transcriptional regulation of RNA polymerase synthesis. *Mol. Cell. Biochem.* 31, 177–196. doi: 10.1007/bf00225850

Ishihama, A., Shimada, T., and Yamazaki, Y. (2016). Transcription profile of *Escherichia coli*: genomic SELEX search for regulatory targets of transcription factors. *Nucleic Acids Res.* 44, 2058–2074. doi: 10.1093/nar/gkw051

Ishihama, A. (1988). Promoter selectivity of prokaryotic RNA polymerase. *Trends Genet.* 4, 282–286. doi: 10.1016/0168-9525(88)90170-9

Ishihama, A. (1990). Molecular assembly and functional modulation of *Escherichia coli* RNA polymerase. *Adv. Biophys.* 26, 19–31. doi: 10.1016/0065-227x(90)90005-e

Ishihama, A. (1997). Adaptation of gene expression in stationary phase bacteria. *Curr. Opin. Genet. Dev.* 7, 582–588. doi: 10.1016/S0959-437X(97)80003-2

Ishihama, A. (1999). Modulation of the nucleoid, the transcription apparatus, and the translation machinery in bacteria for stationary phase survival. *Genes Cells* 3, 135–143. doi: 10.1046/j.1365-2443.1999.00247.x

Ishihama, A. (2000). Functional modulation of *Escherichia coli* RNA polymerase. *Annu. Rev. Microbiol.* 54, 499–518. doi: 10.1146/annurev.micro.54.1.499

Ishihama, A. (2009).*The nucleoid: an overview. EcoSal—Escherichia coli and Salmonella: Cellular and Molecular Biology*. Neidhardt, FC. Washington, DC: American Society for Microbiology Press, 1672–1682.

Ishihama, A. (2010). Prokaryotic genome regulation: multi-factor promoters, multi-target regulators and hierarchic networks. *FEMS Microbial. Rev.* 34, 628–645. doi: 10.1111/j.1574-6976.2010.00227.x

Ishihama, A. (2012). Prokaryotic genome regulation: a revolutionary paradigm. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* 88, 485–508. doi: 10.2183/pjab.88.485

Ishihama, A., Kori, A., Koshio, E., Yamada, K., Maeda, H., Shimada, T., et al. (2014). Intracellular concentrations of 65 species of transcription factors with known regulatory functions in *Escherichia coli*. *J. Bacteriol.* 196, 2718–2727. doi: 10.1128/JB.01579-14

Izutsu, K., Wada, C., Komine, Y., Sako, T., Ueguchi, C., Nakura, S., et al. (2001). *Escherichia coli* ribosome-associated protein SRA, whose copy number increases during stationary phase. *J. Bacteriol.* 183, 2765–2773. doi: 10.1046/j.1365-2443.2001.00457.x

Jaworski, A., Higgins, N. P., Wells, R. D., and Zacarias, W. (1991). Topoisomerase mutants and physiological conditions control supercoiling and Z-DNA formation *in vivo*. *J. Biol. Chem.* 266, 2576–2581.

Jin, D. J., Cagliero, C., and Zhou, Y. N. (2012). Growth rate regulation in *Escherichia coli*. *FEMS Microbiol. Rev.* 36, 269–287. doi: 10.1111/j.1574-6976.2011.00279.x

Jishage, M., and Ishihama, A. (1995). Regulation of RNA polymerase sigma subunit synthesis in *Escherichia coli*: intracellular levels of σ^{70} and σ^{38}. *J. Bacteriol.* 177, 6832–6835. doi: 10.1128/jb.178.18.5447-5451.1996

Jishage, M., and Ishihama, A. (1998). A stationary phase protein in *Escherichia coli* with binding activity to the major sigma subunit of RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 95, 4953–4958. doi: 10.1073/pnas.95.9.4953

Jishage, M., and Ishihama, A. (1999). Transcriptional organization and *in vivo* role of the *Escherichia coli rsd* gene, encoding the regulator of RNA polymerase sigma D. *J. Bacteriol.* 181, 3768–3776.

Jishage, M., Iwata, A., Ueda, S., and Ishihama, A. (1996). Regulation of RNA polymerase sigma subunit synthesis in *Escherichia coli*: Intracellular levels of four species of sigma subunit under various growth conditions. *J. Bacteriol.* 178, 5447–5451. doi: 10.1128/jb.178.18.5447-5451.1996

Jishage, M., Dasgupta, D., and Ishihama, A. (2001). Mapping of the Rsd contact site on the sigma-70 subunit of *Escherichia coli* RNA polymerase. *J. Bacteriol.* 183, 2952–2956. doi: 10.1128/JB.183.9.2952-2956.2001

Kaczanowska, M., and Ryden-Aulin, M. (2007). Ribosome biogenesis and the translation process in *Escherichia coli*. *Microbiol. Mol. Biol. Rev.* 71, 477–494. doi: 10.1128/MMBR.00013-07

Kato, T., Yoshida, H., Miyata, T., Maki, Y., Wada, A., and Namba, K. (2010). Structure of the 100S ribosome in the hibernation stage revealed by electron cryomicroscopy. *Structure* 18, 719–724. doi: 10.1016/j.str.2010.02.017

Kawakami, K., Saitoh, T., and Ishihama, A. (1979). Biosynthesis of RNA polymerase in *Escherichia coli*. IX. Growth-dependent variations in the synthesis rate,

content and distribution of RNA polymerase. *Mol. Gen. Genet.* 174, 107–116. doi: 10.1007/bf00268348

Koch, A. L. (1996). Similarities and differences of individual bacteria within a clone, in *Escherichia coli and Salmonella*. Ed. Neidhardt, F. C. (Washington, DC: American Society for Microbiology Press), 1640–1661.

Kolter, R., Siegele, D. A., and Tormo, A. (1993). The stationary phase of the bacterial life cycle. *Annu. Rev. Microbiol.* 47, 855–874. doi: 10.1146/annurev.mi.47.100193.004231

Kubitschek, H. E., Baldwin, W. W., and Graetzer, R. (1983). Buoyant density constancy during the cell cycle of *Escherichia coli*. *J. Bacteriol.* 155, 1027–1032.

Kusano, S., Ding, Q., Fujita, N., and Ishihama, A. (1996). Promoter selectivity of *Escherichia coli* RNA polymerase Eσ^{70} and Eσ^{38} holoenzymes: effect of DNA supercoiling. *J. Biol. Chem.* 271, 1998–2004. doi: 10.1074/jbc.271.4.1998

Lee, J.-W., Park, Y.-H., and Seok, Y.-J. (2018). Rsd balances (p)ppGpp level by stimulating the hydrolase activity of SpoT during carbon source downshift in. *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 115, E6845–E6854. doi: 10.1073/pnas.1722514115

Lowen, P. C., and Hengge-Aronis, R. (1994). The role of the sigma factor σ^s (*katF*) in bacterial global regulation. *Annu. Rev. Microbiol.* 48, 53–80. doi: 10.1146/annurev.mi.48.100194.000413

Maaloe, O., and Kjeldgaard, N. O. (1966).*Control of macromolecular synthesis*. Inc., New York: W.A. Benjamin.

Maeda, H., Fujita, N., and Ishihama, A. (2000). Competition among seven *Escherichia coli* sigma subunits: relative binding affinities to the core RNA polymerase. *Nucleic Acids Res.* 28, 3497–3503. doi: 10.1093/nar/28.18.3497

Maki, Y., Yoshida, H., and Wada, A. (2000). Two proteins, YfiA and YhbH, associated with resting ribosomes in stationary phase *Escherichia coli*. *Genes Cells* 5, 965–974. doi: 10.1046/j.1365-2443.2000.00389.x

Makinoshima, H., Nishimura, A., and Ishihama, A. (2002). Fractionation of *Escherichia coli* cell populations at different stages during growth transition to stationary phase. *Mol. Microbiol.* 43, 269–279. doi: 10.1046/j.1365-2958.2002.02746.x

Makinoshima, H., Aizawa, S., Hayashi, H., Miki, T., Nishimura, A., and Ishihama, A. (2003). Growth-phase-coupled alterations in cell structure and function of *Escherichia coli*. *J. Bacteriol.* 185, 1338–1345. doi: 10.1128/jb.185.4.1338-1345.2003

Mannik, J., Castillo, D. E., Yang, D., Siopsis, G., and Mannik, J. (2016). The role of MatP, ZapA and ZapB in chromosomal organization and dynamics in *Escherichia coli*. *Nucleic Acids Res.* 44, 1216–1226. doi: 10.1093/nar/gkv1484

Martinez-Antonio, A., Lomnitz, J. G., Sandoval, S., Aldana, M., and Savgeau, M. A. (2012). Regulatory design governing progression of population growth phases in bacteria. *PloS One* 7, e30654. doi: 10.1073/pnas.1722514115

Matzov, D., Bashan, A., Yap, M. F., Amunts, A., and Yonath, A. (2019). Stress response as implemented by hibernating ribosomes: a structural overview. *FEBS J.* 286, 3558–3565. doi: 10.1111/febs.14968

Mehta, P., Jovanovic, G., Ying, L., and Buck, M. (2015). Is the cellular and molecular machinvery in the stationary phase of *Escherichia coli*?. *Biochem. Soc Trans.* 43, 168–171. doi: 10.1042/BST20140267

Mihoub, M., Abdallah, J., Gontero, B., Dairou, J., and Richarme, G. (2015). The DJ-1 superfamily member Hsp31 repairs proteins from glycation by methylglyoxal and glyoxal. *Biochem. Biophys. Res. Commun.* 463, 1305–1310. doi: 10.1016/j.bbrc.2015.06.111

Mitchell, J. E., Oshima, T., Piper, S. E., Webster, C. L., Westblade, L. F., Karimova, G., et al. (2007). The *Escherichia coli* regulator of σ^{70} protein, Rsd, can up-regulate some stress-dependent promoters by sequestering σ^{70}. *J. Bacteriol.* 189, 3489–3495. doi: 10.1128/JB.00019-07

Nachin, L., Nanmark, U., and Nystrom, T. (2005). Differential roles of the universal stress proteins of *Escherichia coli* in oxidative stress resistance, adhesion, and motility. *J. Bacteriol.* 187, 6265–6272. doi: 10.1128/JB.187.18.6265-6272.2005

Niven, G. W. (2004). Ribosome modulation factor protects *Escherichia coli* during heat stress, but this may not be dependent on ribosome dimerization. *Arch. Microbiol.* 182, 60–66. doi: 10.1007/s00203-004-0698-9

Nomura, M., Gource, R., and Baghman, G. (1984). Regulation of the synthesis of ribosomes and ribosomal components. *Annu. Rev. Biochem.* 53, 75–117. doi: 10.1146/annurev.bi.53.070184.000451

Paget, M. S. (2015). Bacterial sigma factors and anti-sigma factors: Structure, function and distribution. *Biomolecules* 5, 1245–1265. doi: 10.3390/biom5031245

Park, Y. H., Lee, C. R., Choe, M., and Seok, Y. J. (2013). HPr antagonizes the anti-s70 activity of Rsd in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 110, 21142–21147. doi: 10.1073/pnas.1316629111

Patikoglou, G. A., Westblade, L. F., Campbell, E. A., Lanour, V., Lane, W. J., and Darst, S. A. (2007). Crystal structure of the *Escherichia coli* regulator of σ^{70}, Rsd, in complex with σ^{70} domain 4. *J. Mol. Biol.* 21, 649–659. doi: 10.1016/j.jmb.2007.06.081

Pletnev, P., Osterman, I., Sergiev, P., Bogdanov, O., and Dontsova, O. (2015). Survival guide: *Escherichia coli* in the stationary phase. *Acta Nat.* 7, 22–33.

Prossliner, T., Skovbo Winther, K., Serensen, M. A., and Gerdes, K. (2018). Ribosome hibernation. *Ann. Rev. Genet.* 52, 321–348. doi: 10.1146/annurev-genet-120215-035130

Puri, P., Eckhardt, T. H., Franken, L. E., Fusetti, F., Stuart, M. C., Boekema, E. J., et al. (2014). Lactococcus lactis YfiA is necessary and sufficient for ribosome dimerization. *Mol. Microbiol.* 91, 394–407. doi: 10.1111/mmi.12468

Raivio, T. L. (2005). Envelope stress responses and Gram-negative bacterial pathogenesis. *Mol. Microb.* 56, 1119–1128. doi: 10.1111/j.1365-2958.2005.04625.x

Richarme, G., Liu, C., Mihoub, M., Abdallah, J., Leger, T., Joly, N., et al. (2017). Gualine glycation repair by DJ-1/Park7 and its bacterial homologs. *Science* 357, 208–211. doi: 10.1126/science.aag1095

Rodionova, I. A., Zhang, Z., Mehla, J., Goodacre, N., Babum, N., Emili, A., et al. (2017). The phosphocarrier protein HPr of the bacterial phosphotransferase system globally regulates energy metabolism by directly interacting with multiple enzymes in*Escherichia coli*. *J. Biol. Chem.* 292, 14250–14257. doi: 10.1074/jbc.M117.795294

Roszak, D. B., and Colwell, R. R. (1987). Survival strategies of bacteria in the natural environment. *Microbiol. Rev.* 51, 365–379.

Saint-Ruf, C., Pesut, J., Sopta, M., and Matic, I. (2007). Causes and consequences of DNA repair activity modulation during stationary phase in *Escherichia coli*. *Crit. Rev. Biochem. Mol. Biol.* 42, 259–270. doi: 10.1080/10409230701495599

Sanchuki, H. B., Gravina, F., Rodrigues, T. E., Gerhardt, E. C., Pedrosa, F. O., Souza, E. M., et al. (2017). Dynamics of the *Escherichia coli* proteome in response to nitrogen starvation and entry into the stationary phase. *Biochim. Biophys. Acta Protetins Proteom.* 1865, 344–352. doi: 10.1016/j.bbapap.2016.12.002

Serra, D. O., and Hengge, R. (2014). Stress responses go three dimensional – the special order of physiological differentiation in bacterial macrocology biofilms. *Environ. Microbiol.* 16, 1455–1471. doi: 10.1111/1462-2920.12483

Shajani, Z., Sykes, M. T., and Williamson, J. R. (2011). Assembly of bacterial ribosomes. *Annu. Rev. Biochem.* 80, 501–526. doi: 10.1146/annurev-biochem-062608-160432

Shimada, T., Makinoshima, H., Ogawa, Y., Miki, T., Maeda, M., and Ishihama, A. (2004). Classification and strength measurement of stationary-phase promoters by use of a newly developed promoter cloning vector. *J. Bacteriol.* 186, 7112–7122. doi: 10.1128/JB.186.21.7112-7122.2004

Shimada, T., Yoshida, H., and Ishihama, A. (2013). Involvement of cyclic AMP receptor protein in regulation of the *rmf* gene encoding the ribosome modulation factor in *Escherichia coli*. *J. Bacteriol.* 195, 2212–2219. doi: 10.1128/JB.02279-12

Shimada, T., Shimada, K., Matsui, M., Kitai, Y., Igarashi, J., Suga, H., et al. (2014). Roles of cell division control factor SdiA: recognition of quorum sensing signals and modulation of transcription regulation targets. *Genes Cells* 19, 405–418. doi: 10.1111/gtc.12139

Stewart., P. S., and Franklin, M. J. (2008). Physiological heterogeneity in biofilms. *Nat. Rev. Microbiol.* 6, 199–210. doi: 10.1038/nrmicro1838

Talukder, A. Z., Iwata, A., Ueda., A., and Ishihama, A. (1999). Growth phase-dependent variation in the protein composition of *Escherichia coli* nucleoid. *J. Bacteriol.* 181, 6361–6370.

Terui, Y., Tabei, Y., Akiyama, M., Higashi, K., Tomitori, H., Yamamoto, K., et al. (2010). Ribosome modulation factor, an important protein for cell viability encoded by the polyamine modulon. *J. Biol. Chem.* 285, 28698–28707. doi: 10.1074/jbc.M110.111195

Trevino-Quintanilla, L. G., Freyre-Gonzalez, J. A., and Martinez-Flores, I. (2013). Anti-sigma factors in *E. coli:* common regulatory mechanisms controlling sigma factors availability. *Curr. Genomics* 14, 378–387. doi: 10.2174/1389202911314060007

Ueta, M., Yoshida, H., Wada, C., Baba, T., Mori, H., and Wada, A. (2005). Ribosome binding proteins YhbH and YfiA have opposite functions during 100S formation in the stationary phase of *Escherichia coli*. *Genes Cells* 10, 1103–1112. doi: 10.1111/j.1365-2443.2005.00903.x

Ueta, M., Ohniwa, R. L., Yoshida, H., Maki, Y., Wada, C., and Wada, A. (2008). Role of HPF (hibernation promoting factor) in translational activity in *Escherichia coli*. *J. Biochem.* 143, 425–433. doi: 10.1093/jb/mvm243

Ueta, M., Wada, C., Daifuku, T., Sako, Y., Bessho, Y., Kitamura, A., et al. (2013). Conservation of two distinct types of 100S ribosome in bacteria. *Genes Cells* 18, 554–574. doi: 10.1111/gtc.12057

Vogel, H. J., and Bonner, D. M. (1956). Acetylornithinase of *Escherichia coli*: partical purification and some properties. *J. Biol. Chem*. 218, 97–106.

Wada, A., and Sako, T. (1987). Primary structures of and genes for new ribosomal proteins A and B in *Escherichia coli*. *J. Biochem.* 101, 817–820. doi: 10.1093/jb/101.3.817

Wada, A., Yamazaki, Y., Fujita, N., and Ishihama, A. (1990). Structure and probable genetic location of a "ribosome modulation factor" associated with 100S ribosome in stationary-phase *Escherichia coli* cells. *Proc. Natl. Acad. Sci. U.S.A.* 87, 2657–2661. doi: 10.1073/pnas.87.7.2657

Wada, A., Igarashi, K., Yoshimura, S., Aimoto, S., and Ishihama, A. (1995). Ribosome modulation factor: stationary growth phase-specific inhibitor of ribosome functions from *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 214, 410–417. doi: 10.1006/bbrc.1995.2302

Wada, A. (1986a). Analysis of *Escherichia coli* ribosomal proteins by an improved two dimensional gel electrophoresis. I. Detection of four new proteins. *J. Biochem. (Tokyo)* 100, 1583–1594. doi: 10.1093/oxfordjournals.jbchem.a121866

Wada, A. (1986b). Analysis of *Escherichia coli* ribosomal proteins by an improved two dimensional gel electrophoresis. II. Characterization of four new proteins. *J. Biochem. (Tokyo)* 100, 1595–1605. doi: 10.1093/oxfordjournals.jbchem.a121867

Wada, A. (1998). Growth phase coupled modulation of *Escherichia coli* ribosomes. *Genes Cells* 3, 203–208. doi: 10.1046/j.1365-2443.1998.00187.x

Westblade, L. F., Ilag, L. L., Powell, A. K., Kolb, A., Robinson, C. V., and Busby, S. J. W. (2004). Studies of the Escherichia coli Rsd-sigma 70 complex. *J. Mol. Biol.* 335, 685–692. doi: 10.1016/j.jmb.2003.11.004

Yamagishi, M., Matsushima, H., Wada, A., Sakagami, M., Fujita, N., and Ishihama, A. (1993). Regulation of the *Escherichia coli rmf* gene encoding ribosome modulation factor (RMF): Growth phase- and growth rate-control. *EMBO J.* 12, 625–630. doi: 10.1002/j.1460-2075.1993.tb05695.x

Yamamoto, N., Nakahigashi, K., Nakamichi, T., Yoshino, M., Takai, Y., Touda, Y., et al. (2009). Update on the Keio collection of *Escherichia coli* single-gene deletion mutants. *Mol. Syst. Biol.* 5, 335. doi: 10.1038/msb.2009.92

Yoshida, H., and Wada, A. (2014). The 100S ribosome: ribosomal hibernation induced by stress. *Wiley Interdiscip. Rev. RNA* 2014, 723–732. doi: 10.1002/wrna.1242

Yoshida, H., Maki, Y., Kato, H., Fujisawa, H., Izutsu, K., Wada, C., et al. (2002). The ribosome modulation factor (RMF) binding site on the 100S ribosome of *Escherichia coli*. *J. Biochem.* 132, 983–989. doi: 10.1093/oxfordjournals.jbchema003313.

Yoshida, H., Yamamoto, H., Uchiumi, T., and Wada A. (2004). RMF inactivates ribosomes by covering the peptidyl transferase centre and entrance of peptide exit tunnel. *Genes Cells* 9, 271–278. doi: 10.1111/j.1356-9597.2004.00723.x

Yoshida, H., Shimada, T., and Ishihama, A. (2018). Coordinated hibernation of transcriptional and translational apparatus expression during growth transition of *Escherichia coli* into stationary phase. *mSystems* 3, e00057–e00018. doi: 10.1128/mSystems 00057-18.

Yuan, A. H., Gregory, B. D., Sharp, J. S., McCleary, K. D., Dove, S. L., and Hochschild, A. (2008). Rsd family proteins make simultaneous interactions with regions 2 and 4 of the primary sigma factor. *Mol. Microbiol.* 70, 1136–1151. doi: 10.1111/j.1365-2958.2008.06462.x

Zengel, J. M., and Lindahl, L. (1994). Diverse mechanisms for regulating ribosomal protein synthesis in *Escherichia coli*. *Prog. Nucleic Acid Res. Mol. Biol.* 47, 331–370. doi: 10.1016/s0079-6603(08)60256-1

Zinser, E. R., and Kolter, R. (2004). *Escherichia coli* evolution during stationary phase. *Res. Microbiol.* 155, 328–336. doi: 10.1016/j.resmic.2004.01.014

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

ORIGINAL RESEARCH
published: 14 February 2020
doi: 10.3389/fgene.2019.01356

The Sole *Mycobacterium smegmatis* MazF Toxin Targets $tRNA^{Lys}$ to Impart Highly Selective, Codon-Dependent Proteome Reprogramming

Valdir Cristovao Barth[1] *and Nancy A. Woychik*[1,2]*

[1] Department of Biochemistry and Molecular Biology, Robert Wood Johnson Medical School, Rutgers University, Piscataway, NJ, United States, [2] Rutgers Cancer Institute of New Jersey, New Brunswick, NJ, United States

Survival of mycobacteria, both free-living and host-dependent pathogenic species, is dependent on their ability to evade being killed by the stresses they routinely encounter. Toxin-antitoxin (TA) systems are unique to bacteria and archaea and are thought to function as stress survival proteins. Here, we study the activity of the endoribonuclease toxin derived from the MazEF TA system in *Mycobacterium smegmatis*, designated MazEF-ms. We first enlisted a specialized RNA-seq method, 5' RNA-seq, to identify the primary RNA target(s) of the MazF-ms toxin. Just two tRNA species, $tRNA^{Lys\text{-}UUU}$ and $tRNA^{Lys\text{-}CUU}$, were targeted for cleavage by MazF-ms at a single site within their anticodon sequence (UU↓U and CU↓U) to render these tRNAs nonfunctional for protein synthesis. The 5' RNA-seq dataset also revealed hallmarks of ribosome stalling predominantly at Lys AAA codons even though both Lys tRNAs were cleaved by MazF-ms. Stalled ribosomes were then cleaved on their 5' side by one or more RNases, resulting in very selective degradation of only those mRNAs harboring ribosomes stalled at Lys codons. This highly surgical, codon-dependent degradation of mRNA transcripts was validated using quantitative mass spectrometry of proteins that were newly synthesized during MazF-ms expression. The *M. smegmatis* proteome was altered as predicted, Lys AAA codon-rich proteins was downregulated while Lys AAA codon deficient proteins were upregulated. Analysis of specific subsets of proteins that were upregulated or downregulated was consistent with the growth-arrested phenotype of MazF-ms expressing cells. Curiously, the tRNA target and mechanism of action of MazF-ms paralleled that of one atypical MazF toxin in *M. tuberculosis*, suggesting manipulation of the levels of lysine tRNAs as the preferred conduit for reprogramming the proteomes *via* ribosome stalling at rare AAA codons in these GC-rich mycobacteria.

Keywords: MazF, toxin, antitoxin, transcriptome, ribosome, translation

OPEN ACCESS

Edited by:
Philipp Kapranov,
Huaqiao University, China

Reviewed by:
Thomas Begley,
University at Albany, United States
Xiaoyun Wang,
South China Normal University, China

***Correspondence:**
Nancy A. Woychik
nancy.woychik@rutgers.edu

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *10 September 2019*
Accepted: *11 December 2019*
Published: *14 February 2020*

Citation:
Barth VC and Woychik NA (2020) The Sole Mycobacterium smegmatis MazF Toxin Targets $tRNA^{Lys}$ to Impart Highly Selective, Codon-Dependent Proteome Reprogramming. Front. Genet. 10:1356. doi: 10.3389/fgene.2019.01356

INTRODUCTION

Stress is a constant threat to the survival of free-living organisms. Toxin-antitoxin (TA) systems are believed to act as one line of defense enlisted by free-living bacteria to survive the constant barrage of environmental assaults in their native habitats (reviewed in Yamaguchi et al., 2011; Harms et al., 2018). TA systems are operons comprising adjacent genes encoding two small (~10 kDa) proteins, an antitoxin and its cognate toxin. Based on current models developed through studies of *Escherichia coli* TA systems (reviewed in Yamaguchi et al., 2011; Harms et al., 2018), in the absence of stress the intrinsic activity of the toxin is sequestered by the formation of a stable TA protein-protein complex. However, in response to one or more specific stress triggers, the antitoxin is degraded by proteases. The resulting paucity of antitoxin results in an excess of free toxin which then acts on its specific intracellular target. Toxin action on its target(s) generally results in growth arrest which is characteristically reversible when the stress is released, enabling replenishment of antitoxin. Thus, TA systems appear to work well for pulses of stress instead of prolonged stress where a major physiological and structural transformation is warranted, i.e., conversion from vegetative state to a nearly dormant spore.

Mycobacterium smegmatis is most recognized as a laboratory surrogate for the study of features it shares with its pathogenic relative, *M. tuberculosis*. However, in nature *M. smegmatis* is a saprophyte that lives in ever changing environments within soil, water and on plants. Thus, it should be genetically hardwired for stress survival even though its genome harbors only a small fraction of TA systems compared to *M. tuberculosis* [three in *M. smegmatis* (Frampton et al., 2012) vs. ~90 in *M. tuberculosis* (Ramage et al., 2009; Sala et al., 2014)]. In this study, we determine the intracellular target and study the function of one of the three distinct TA system family toxins in *M. smegmatis*, the sole MazF family member that we designate as MazF-ms (MSMEG_4448). We found that the primary RNA targeted for cleavage by this endoribonuclease toxin was tRNA^{LysUUU}, identical to one of the 11 MazF toxins in *M. tuberculosis,* MazF-mt9 (Schifano et al., 2016; Barth et al., 2019). We then document a spectrum of downstream events that lead to surgical remodeling of the proteome, which closely paralleled that which occur in *M. tuberculosis* (Barth et al., 2019). This striking conservation of tRNA targets underscores the importance of the relatively rare Lys AAA mRNA codon as an efficient conduit for modulating the physiology in both mycobacterial species through activation of MazF toxins.

MATERIALS AND METHODS

Strains, Plasmids, and Reagents

All experiments were performed using *Mycobacterium smegmatis* strain mc^2 155 (ATCC 700084). *M. smegmatis* cells were grown at 37°C in Difco Middlebrook 7H9 media (BD) supplemented with 5 g/L albumin, 2 g/L dextrose, 0.085 g/L NaCl, 0.05% Tween 80 and 25 μg/ml kanamycin (for plasmid selection), under constant shaking at 200 rpm.

The gene MSMEG_4448, here referred as MazF-ms, was amplified by PCR from *M. smegmatis* genomic DNA using the oligos NWO2791 (5'-AGA TAC ATA TGC GGC GCG GCG ATA TCT ACA CCG CGG-3') and NWO2792 (5'-AGA TAA AGC TTC ACC CGG CGA TTC CCA GAA AAA CC-3'). The amplified DNA was cloned into the anhydrotetracycline (ATC)-inducible plasmid pMC1s (Ehrt et al., 2005) modified to substitute unique NdeI-HindIII sites in place of ClaI-EcoRI to enable insertion of a gene with 5'NdeI-3'HindIII sites. MazF-ms expression was induced by adding ATC to the media at a final concentration of 200 ng/ml when cells reached an OD (600 nm) between 0.1 and 0.2 and compared to uninduced (-ATC) samples.

RNA Isolation

In order to extract total RNA, ~50 ml of *M. smegmatis* cells were collected by centrifugation at 2000 g at 4°C for 5 min. Cell pellets were resuspended in Tri reagent (Zymo Research) and transferred to 2 ml lysing kit tubes (Bertin Corp.) containing 0.1 mm glass beads. Cells lysis was performed on a Precellys Evolution homogenizer (Bertin Corp.) by three consecutive 30-s pulses at 9,000 rpm, with 1 min cooling periods on ice in between each cycle. The samples were centrifuged for 5 min at 14,000 rpm at 4°C, and RNA was isolated from the supernatant using the Direct-zol RNA Miniprep Plus extraction kit (Zymo Research). After isolation, the samples were treated with 1 U of Turbo DNase as an extra genomic DNA removal step, purified using the RNA Clean and Concentrator kit (Zymo Research) and eluted in 40 μl of RNase-free water. RNA concentration was measured in a BioSpectrometer (Eppendorf) with a μCuvette.

5′ RNA-Seq

5' OH libraries were constructed as previously described (Schifano et al., 2014). Briefly, in order to remove 5' monophosphate RNA species, three μg of the purified RNA from induced and uninduced cultures were digested with 1 U Terminator at 30°C for 1 h. After purification using the RNA Clean and Concentrator kit (Zymo Research), the samples were phosphorylated using 3 U of T4 PNK at 37°C for 1 h and re-purified with the same kit. 5' adapter (5′-GUUCAGAGU UCUACAGUCCGACGAUCNNNNNN-3′) was ligated using 1 U of T4 RNA ligase 1 (New England Biolabs) at 16°C for approximately 18 h. In order to remove the remaining free adapters, the adapter-ligated RNAs were resolved on a 6% TBE-Urea PAGE gel, excised and precipitated in isopropanol at −20°C. The purified RNAs were used in a reverse transcription reaction using Superscript IV (Thermo Fisher) and the degenerate primer (5′-GCCTTGGCACCCGAGAATTCCA NNNNNNNNN-3′). The resulting cDNA was loaded into a 10% TBE-Urea gel and fragments between 80 and 500 nts were excised and precipitated. The cDNA libraries were amplified in a PCR reaction with Phusion HF DNA Polymerase (Thermo Fisher). The primers used were RP1 (5′- AATGA TACGGCGACCACCGAGATCTACACGTTCAGAGTTCTA CAGTCCGA -3′) and RPIX (5′-CAAGCAGAAGACGGCATA CGAGATNNNNNNGTGACTGGAGTTCCTTGG CACCCGAGAATTCCA-3′), where the N's represent the

individual Illumina barcodes for each library. The amplified libraries between 150–450 bp were gel purified and subjected to single-end sequencing in an Illumina HiSeq 2500 or HiSeq4000 sequencer.

Note that our 5' RNA-seq method specifically selects for RNA molecules with 5' hydroxyl ends created upon cleavage by MazF-ms. Based on the detailed schematic of our method in Schifano et al. (2016), if reverse transcriptase should pause at tRNA modifications, the truncated cDNA would not contain the complement to the 5' adapter sequence that was exclusively ligated to RNAs containing a 5'-OH. Without the adapter sequence, truncated cDNAs would not be amplified by PCR nor could they be sequenced by the Illumina primer which is also complementary to the 5' adapter.

The resulting FASTQ files had the adapter sequences and the first 6 nucleotides of the 5' end trimmed using Trimmomatic (Bolger et al., 2014). Reads were trimmed to 20 nts and the ones containing fewer than 20 nucleotides were excluded. The remaining reads were mapped to *M. smegmatis* genome (NCBI accession: CP000480.1) using bowtie 1.2 applying the parameters –n 0–l20 (Langmead et al., 2009). Next, we calculated the number of reads that started at a given genome position for each nucleotide in the genome. Genomic positions with 0 counts received a pseudo count of 1 in the uninduced sample. The counts were normalized by sequencing depth, in reads per million (rpm) of mapped reads and the counts of the induced sample were divided by the uninduced control to generate a fold change. Unless otherwise stated, we only considered positions with at least 50 rpm and 5 rpm in the induced sample for tRNA and mRNA genes, respectively, and a fold change >10. Sequence and frequency logos were generated by kpLogo (Wu and Bartel, 2017). The fastq files were submitted to NCBI's Sequence Read Archive (SRA), under BioProject number PRJNA564437.

Labeling of Newly Synthesized Proteins and Proteomic Analysis

To identify and quantify which proteins are translated after induction of MazF-ms (MSMEG_4448), three biological replicates were grown to an OD_{600} between 0.1 and 0.2 and divided into induced (+ATC) and uninduced samples. In order to label newly synthesized proteins, the methionine analog azidohomoalanine (AHA, AnaSpec) was added to the media at 50 μM after 4.5 h of MazF-ms induction. After 1.5 h of incubation with AHA, the cells were pelleted by centrifugation at 2000 g at 4°C for 10 min and washed with PBS. The cells were resuspended in a 2% CHAPS, 8M Urea buffer and lysed as described in the "RNA isolation" section, using Precellys Evolution homogenizer (Bertin Corp). The cell lysate was centrifuged and the AHA-containing proteins in the supernatant were captured using Click-iT™ Protein Enrichment Kit (ThermoFisher) followed by in-column trypsin digestion.

Digests were analyzed in two separate runs and combined. The data was analyzed as described previously (Barth et al., 2019), considering proteins with ≥15 detected spectral counts. Q-values are calculated using the fdrtool package of Strimmer (Strimmer, 2008) with significant changes at or below a q-value of 0.05. The raw files were deposited in Mass Spectrometry Interactive Virtual Environment (MassIVE) repository (accession number: MSV000084300).

RESULTS

MazF-ms Expression Arrests Growth in *M. smegmatis*

According to the Toxin-Antitoxin Database (TADB 2.0) (Xie et al., 2018), the *M. smegmatis* reference genome harbors only one gene from the MazF toxin family, annotated as MSMEG_4448 (here referred to as MazF-ms). To establish whether or not MazF-ms is toxic (i.e., leads to cell growth arrest) when expressed in *M. smegmatis*, we cloned the gene under the control of an ATC-inducible promoter in the pMC1s plasmid and transformed *M. smegmatis* mc^2 155 cells. The expression level from this plasmid is modest, a 4.5-fold induction, based on measurement of mCherry fluorescence (**Supplementary Figure 1**). MazF-ms expression led to pronounced growth arrest that started between 4.5 to 6 h (**Figure 1**) and was sustained for at least 18 h.

MazF-ms Exclusively Targets Both tRNA^{Lys} Isoacceptors for Cleavage at a Single Site Within Their Anticodons

All members of the MazF family reported to date are single-strand, sequence-specific endoribonucleases. To help elucidate the molecular mechanism by which the ribonuclease activity of MazF-ms is able to regulate growth, we applied our specialized RNA seq method, 5' RNA-seq, to find its RNA target(s)

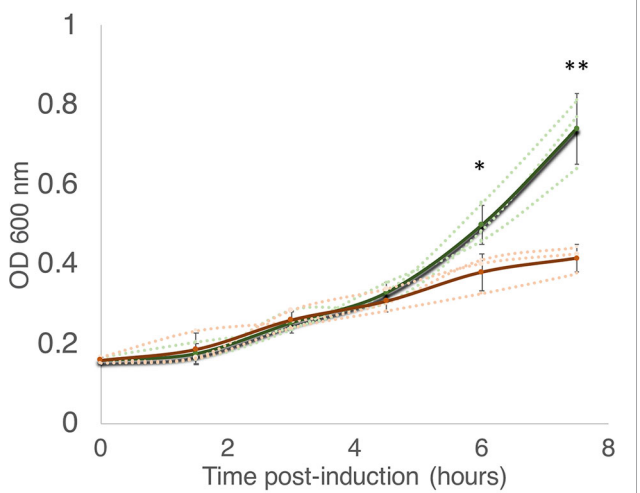

FIGURE 1 | MazF-ms expression in *M. smegmatis* leads to growth inhibition. *M. smegmatis* cells harboring pMC1s-MazF-ms were grown in triplicate in supplemented 7H9 media until OD_{600nm} = ~ 0.16. The cultures were split into induced (+ATC, orange lines) and uninduced (-ATC, green lines) and absorbance at 600 nm was determined every 1.5 h. Error bars represent the standard deviation from the average (solid line) of the three biological replicates (dotted lines). Asterisks represent statistical significance between induced and uninduced in a Student's t test comparison (*, p-value = 0.011; **, p-value = 0.0046).

(Schifano et al., 2014). This technique was originally developed to selectively sequence transcripts based on their 5' ends. Here, we apply 5' RNA-seq to identify RNA fragments containing a 5' hydroxyl (5' OH) end, which are products of MazF toxin activity (Schifano et al., 2014). This approach also allows us to precisely map the cleavage position at a single nucleotide resolution.

Given that the difference in growth between induced and uninduced cultures is more dramatically observed after 6 h post induction, we selected two time points for RNA isolation: one immediately before we observed growth separation (at 4.5 h) and one where the separation is significant (at 6 h). In both time points, 5' RNA-seq identified internal cleavage of the only two tRNA^{Lys} isoacceptors [tRNA$^{Lys23\text{-}UUU}$ and tRNA$^{Lys18\text{-}CUU}$ (Lowe and Chan, 2016)] annotated in the *M. smegmatis* genome (**Figures 2A–D**). When compared to controls, MazF-ms-induced datasets showed an enrichment of 41 to 163-fold in intragenic 5' OH ends at position 36 of both tRNA^{Lys} genes. There are 46 standard tRNAs in *M. smegmatis* and one

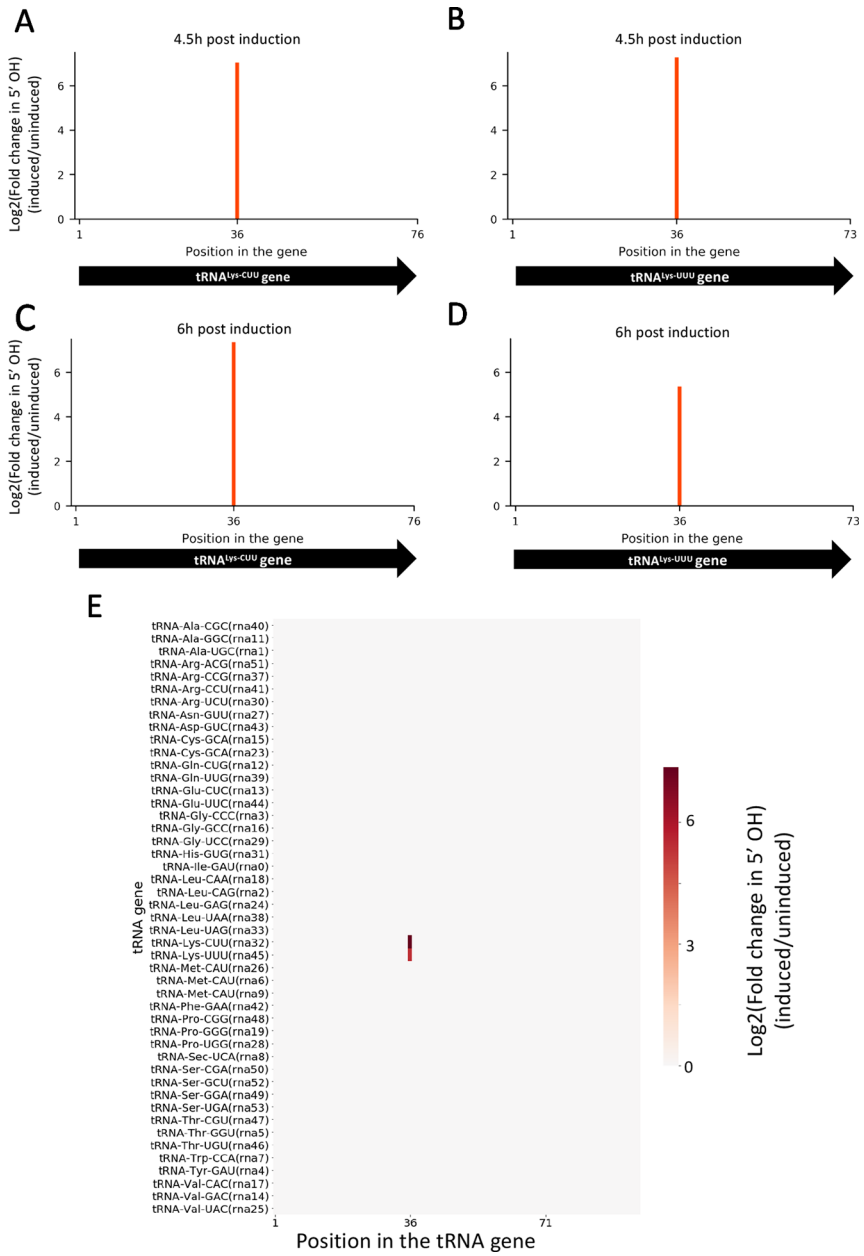

FIGURE 2 | MazF-ms selectively targets both tRNA^{Lys} isoacceptors. **(A–D)** Fold changes of 5'OH (indicating endonucleolytic cleavage) within both tRNA$^{Lys\text{-}CUU}$ and tRNA$^{Lys\text{-}UUU}$ genes detected by 5' RNA-seq. *M. smegmatis* cells expressing MazF-ms for 4.5 **(A, B)** or 6 h **(C, D)** were compared to uninduced controls. **(E)** Heatmap showing the fold change in 5' OH levels in each position of all 47 *M. smegmatis* tRNA genes after 6 h of MazF-ms induction. The annotated tRNA gene ID (from genome CP000480.1) is shown in parentheses.

selenocysteine tRNA, none of the other 45 tRNAs were cleaved by MazF-ms (**Figure 2E**).

Other tRNA-cleaving toxins, such as MazF-mt9 and VapC-mt11 (Schifano et al., 2016; Cintrón et al., 2019), rely on both sequence and secondary structure to accurately recognize their targets. Accordingly, the two tRNA^{Lys} identified here as MazF-ms targets only differ by one nucleotide in the ± 5 nt region surrounding the cleavage site (**Figure 3A**). This site is located in the anticodon stem loop, where the predicted secondary structure is highly conserved. More specifically, cleavage occurred between the second and third bases of the anticodons (UU↓U and CU↓U, **Figure 3B**), presumably inactivating these tRNAs.

5' RNA-Seq of MazF-ms-Expressing Cells Reveals Ribosome Stalling

Having established that the primary targets were tRNA$^{Lys\text{-}UUU}$ and tRNA$^{Lys\text{-}CUU}$, we questioned whether the depletion of these tRNAs would lead to ribosome stalling in *M. smegmatis* at the mRNA codons requiring these tRNAs as we had previously observed for *M. tuberculosis* (Barth et al., 2019). Indeed, we identified 130 cleaved mRNAs in our 5' RNA seq dataset that were not similar in sequence to the tRNA targets at the cleavage site and are not expected to have the secondary structure requirements for MazF-ms recognition demonstrated by Schifano et al. (2016). When aligned by their 5' ends, these transcripts showed a clear AAA or AAG consensus sequence approximately 15 nt downstream of the 5' OH end (**Figures 4A, B**), the cognate Lys codons for tRNA$^{Lys\text{-}UUU}$ and tRNA$^{Lys\text{-}CUU}$, respectively. As we had recently proven by Ribo-seq, the 15-nt spacing from the codon to the 5' OH end indicates that a stalled ribosome was bound at this position of the mRNA *in vivo* (Barth et al., 2019). The 15 nts represents the approximate distance from the 5' side of the translating ribosome to the A-site (**Figures 4D**, Barth et al., 2019). Thus, evidence of the stalled ribosome is fortuitously revealed by 5' RNA-seq because upon stalling, the mRNA is then cleaved on its 5' side of the ribosome by one or more cellular RNase(s)—not MazF-ms—that generates a 5' OH upon cleavage.

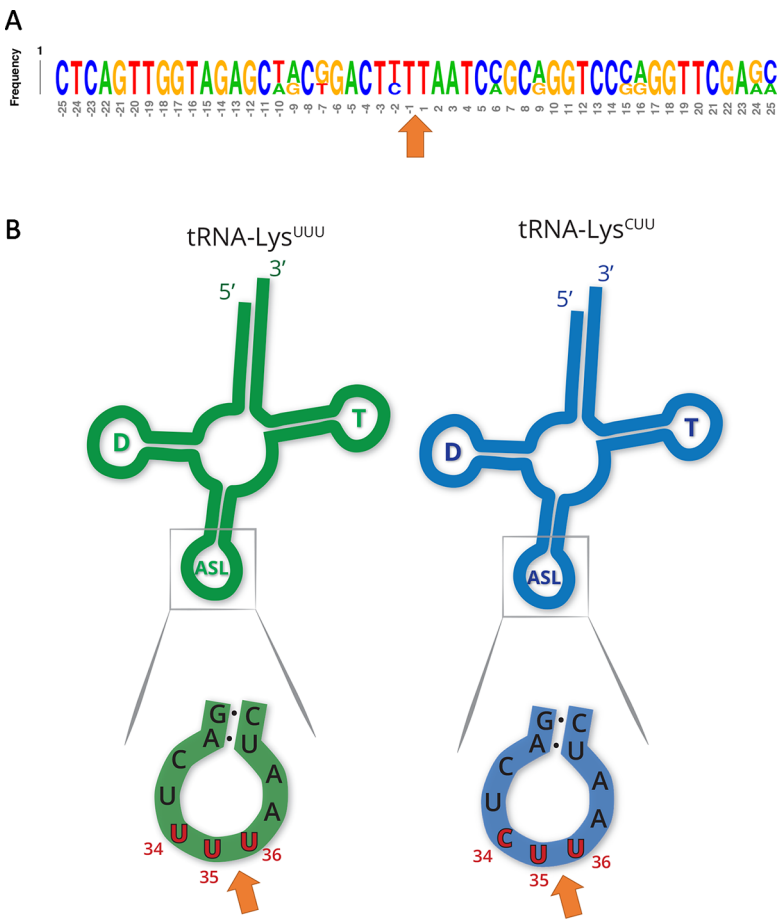

FIGURE 3 | MazF-ms targets show high secondary structure and sequence similarity. **(A)** Frequency logo showing the DNA sequence similarity between tRNA$^{Lys\text{-}UUU}$ and tRNA$^{Lys\text{-}CUU}$ genes in the 50 nucleotides surrounding MazF-ms cleavage position (orange arrow). **(B)** Schematic representation of tRNA$^{Lys\text{-}UUU}$ and tRNA$^{Lys\text{-}CUU}$, illustrating the D-arm (D), T-arm (T) and anticodon stem loop (ASL) portions of the tRNA. ASL is partially shown in greater detail to emphasize the sequence and secondary structure near the cleavage site (orange arrow). Numbering in red indicates the nucleotide position of the anticodon in the mature tRNA molecule.

We have described the same cascade of events for just one of the 11 MazF family members in *M. tuberculosis*, MazF-mt9, in which tRNA$^{Lys43\text{-}UUU}$ depletion leads to ribosome stalling and cleavage on the 5' end of the ribosome. The observation of an analogous trend in the MazF-ms 5' RNA-seq datasets indicating ribosome stalling and subsequent cleavage strongly suggests mechanistic conservation between MazF-mt9 and MazF-ms toxins regarding initial toxin-mediated tRNA cleavage followed by a secondary ribosome stalling/mRNA cleavage event. In *M. smegmatis*, however, although tRNA cleavage was significant for both isoacceptors, the vast majority of the observed stalled ribosomes (75%) paused at the rarer Lys AAA codon rather than the more frequent Lys AAG (**Figure 4C**). Therefore, as in *M. tuberculosis*, our data support a model in which MazF-ms acts by depleting the cellular pool of tRNA^{Lys} causing ribosome stalling at Lys codons (predominantly Lys AAA), followed by

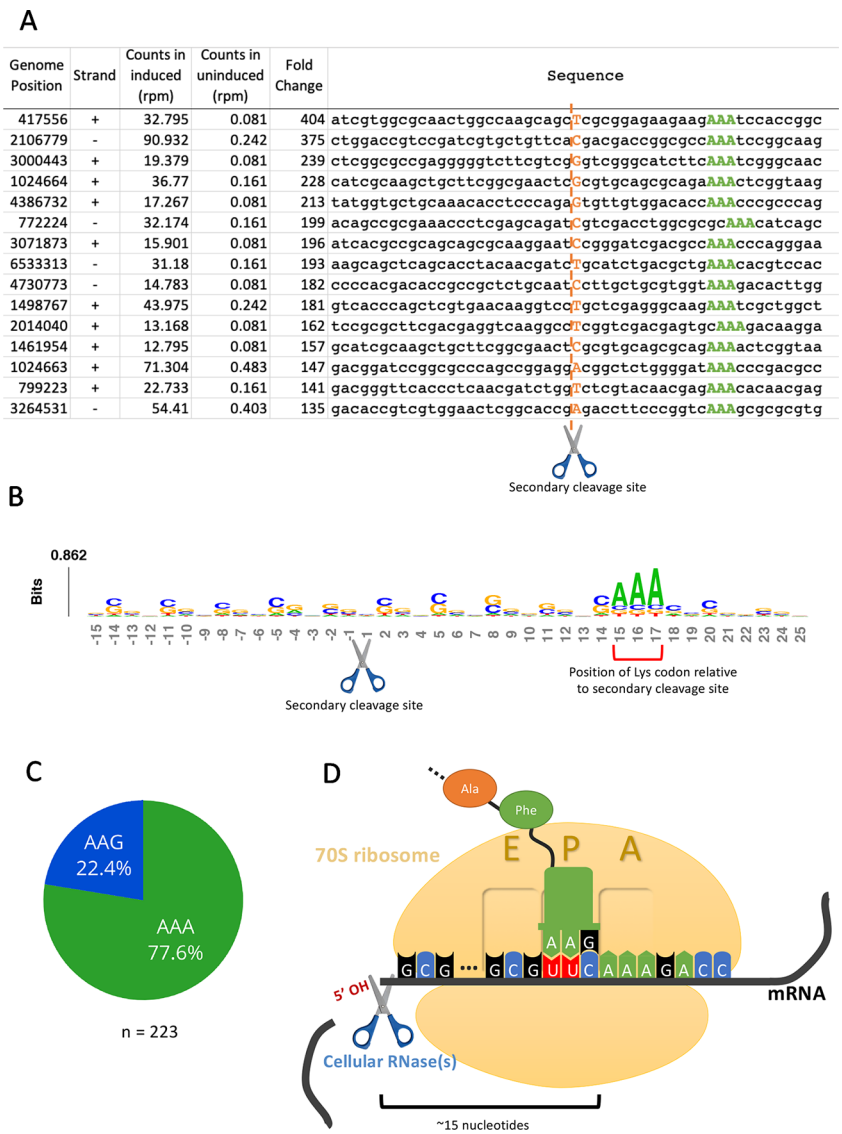

FIGURE 4 | 5'RNA-seq serendipitously reveals ribosome stalling at lysine codons. **(A)** Top mRNA hits found in the 5'RNA-seq dataset. The 50 nucleotides surrounding the secondary cleavage site (dotted line) generated by one or more cellular RNases (scissor) are shown. Lysine AAA codons are indicated in green, approximately 15 nucleotides downstream of the cleavage site. The first nucleotide (adjacent to the 5' OH) of the read is highlighted in orange. Counts are normalized to reads per million (rpm). **(B)** Sequence logo summarizing the 130 detected mRNA hits. The height of each nucleotide is proportional to its frequency at that given position. Positions are numbered relative to the secondary RNase cut site (scissor). Red bracket indicates the position of lysine AAA codons. **(C)** Proportion of stalling in AAA codons vs. AAG codons found by 5' RNA-seq in 223 annotated transcripts with at least 1 rpm in the induced sample and containing a lysine codon at or near +15. **(D)** Schematic representation demonstrating the events following the depletion of the cellular levels of tRNA^{Lys} by MazF-ms. Due to the lack of available tRNA^{Lys}, translating ribosomes stall mainly at lysine AAA codons at the A site. Ribosome stalling events likely trigger mRNA cleavage at 5' side of the stalled ribosome by one or more cellular RNases (scissor).

recruitment of another RNase that cleaves 5' of the stalled ribosome (**Figure 4D**).

MazF-ms Promotes Codon-Specific Translation

Next, we sought to characterize the proteomic changes promoted by the Lys AAA/AAG-specific ribosome stalling events mediated by MazF-ms. In order to distinguish proteins that were only synthesized after MazF-ms induction from preexisting "old" proteins, we adopted a method that utilizes a methionine mimetic called azidohomoalanine (AHA). AHA is incorporated into nascent peptides during translation, therefore only marking proteins that were synthesized after its addition to the media. Due to its azide moiety, proteins containing AHA residues can be captured using an alkyne-containing resin through a Cu(I)-catalyzed click reaction and then analyzed by quantitative mass spectrometry. We added AHA to the cultures after 4.5 h of MazF-ms induction, tagging only proteins that were newly translated.

In contrast to a translation shut-off model proposed for other tRNA-cleaving toxins (Winther et al., 2016), global translation was not halted. One hundred twenty six proteins were significantly more abundant in cells expressing MazF-ms compared to the controls (**Supplementary Table 1**, q-value < 0.05). Interestingly, a striking difference in AAA codon content was detected in these proteins compared to the less abundant ones (**Figure 5A**), i.e. upregulated proteins were generally Lys AAA codon deficient while downregulated proteins were generally Lys AAA codon rich. This trend was not as striking when the proteomics data was instead graphed based on AAG codon content (**Figure 5B**). These trends observed in newly synthesized proteins were concordant with the relative abundance of ribosome stalling events detected in our 5' RNA-seq datasets at Lys AAA codons (>75% of transcripts with evidence of ribosome stalling). Finally, the ability of MazF-ms to preferentially influence the overall cellular pool of Lys AAA codon-containing proteins over those containing Lys AAG codons is also graphed in **Figure 6**.

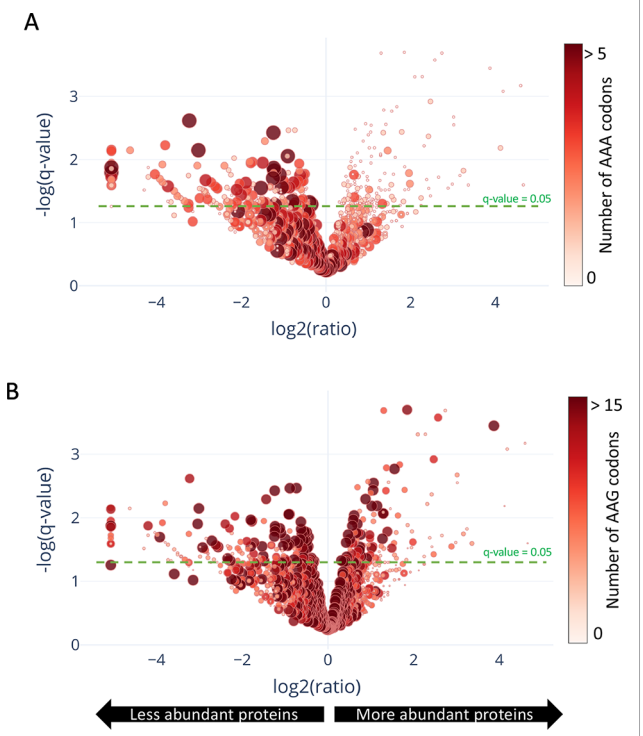

FIGURE 5 | Translation of proteins rich in lysine AAA codons is reduced upon MazF-ms expression. **(A–B)** Volcano plot showing differentially translated proteins (circles) detected by quantitative mass spectrometry and their AAA **(A)** or AAG **(B)** codon content. *M. smegmatis* cultures expressing MazF-ms for 4.5 h were incubated with azidohomoalanine (AHA) to label only newly synthesized proteins after toxin expression. The color saturation and circle size are proportional to the number of AAA **(A)** or AAG **(B)** codons in the gene.

MazF-ms Reduces Translation of Critical Components of the DNA Replication Machinery While Concomitantly Supporting Synthesis of Stress Response Proteins

In TA systems, specific stresses trigger degradation of the cognate antitoxin by a protease, freeing toxin to act within the cell (Yamaguchi et al., 2011; Harms et al., 2018; Song and Wood, 2018). Therefore, overexpression of MazF-ms (used in this analysis and throughout this work) is intended to mimic natural toxin activation that occurs when cells are exposed to the relevant stress. After observing hundreds of differentially translated proteins in **Figure 5**, we analyzed the two datasets comprising the more abundant or less abundant proteins following MazF-ms toxin expression using the Function Annotation Tool in the Database for Annotation, Visualization and Integrated Discovery (DAVID) platform (Huang et al., 2009a; Huang et al., 2009b). This tool within DAVID was used to identify functionally similar proteins which were enriched in the two datasets (**Figure 7**). Since toxin overexpression simulates toxin activation by exposure to stress, the proteins identified using DAVID are expected to be physiologically relevant.

Among the 171 statistically significant downregulated proteins from **Figure 5A** (in blue in **Figure 7**), there were some notable trends consistent with cells in a state of growth arrest (**Supplementary Table 2**). First, there were 11 downregulated enzymes involved in "fatty acid catabolism" (FAD heading; blue box **Figure 7**). Fatty acid catabolism occurs predominantly through successive rounds of β-oxidation, a process whereby even-chain fatty acids are degraded to acetyl-CoA and odd-chain fatty acids are degraded to acetyl-CoA and propionyl-CoA. Second, there were five functionally related proteins within the "DNA replication" and "nucleotide binding" categories (blue box **Figure 7**). In the DNA replication category, one critical protein was downregulated, the only catalytic subunit of the three subunit core DNA polymerase III enzyme (α subunit, MSMEG_3178). This core enzyme is a component of the DNA polymerase III holoenzyme that mediates DNA replication in bacteria. Within the nucleotide binding group were a spectrum of proteins whose downregulation is also logical for cells in a state of

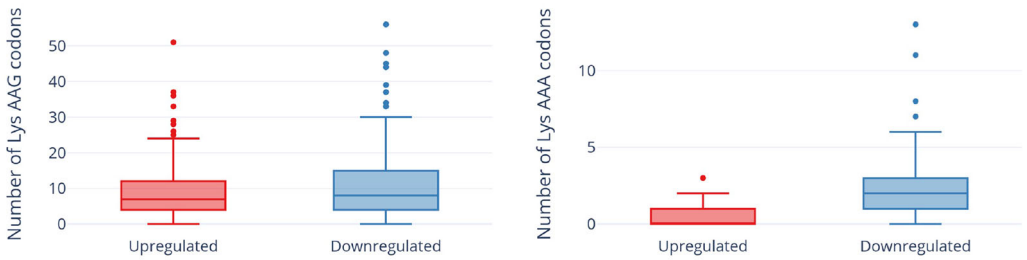

FIGURE 6 | Expression of MazF-ms leads to global proteomic shifts based on the Lys AAA, but not Lys AAG, codon content. Distribution of AAG (left) or AAA (right) codons in significantly upregulated or downregulated proteins. Outliers are shown as individual dots.

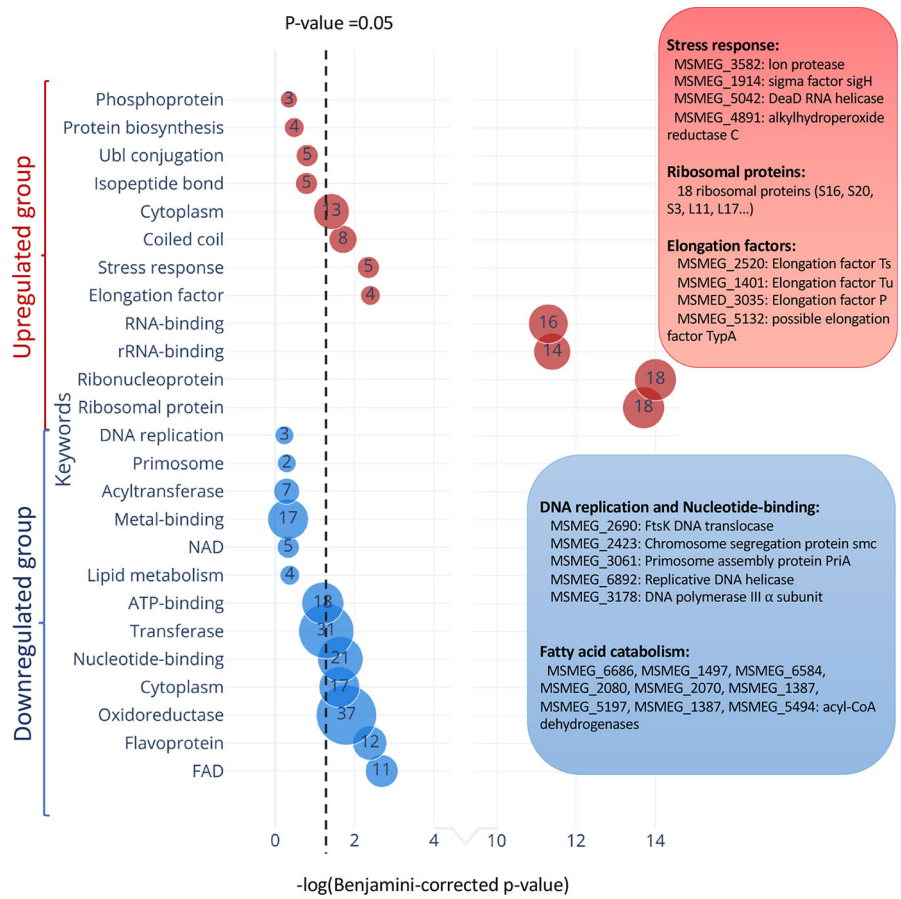

FIGURE 7 | Functional classification of differentially translated proteins during MazF-ms expression. Enriched DAVID UP_KEYWORDS (corresponding to UniProt keywords) were generated by analyzing the upregulated (red) or downregulated (blue) proteins identified in azidohomoalanine (AHA)-proteomics. The diameter of each circle and its numbering correspond to the number of genes associated with the given keyword. Examples of genes in the main groups are described by relevant categories in the corresponding color-matched boxes.

growth arrest: RecA (MSMEG_2723) is involved in DNA repair, the DNA translocase FtsK (MSMEG_2690) is localized at the septum where cell division occurs, chromosome segregation protein SMC (MSMEG_2423) and the priA (MSMEG_3061) are components of the primosome protein complex that activates DNA replication forks.

There were 126 upregulated proteins upon MazF-ms expression. When this dataset was subjected to the Functional Annotation Tool in DAVID, there were several proteins whose upregulation were also consistent with the growth arrested state of MazF-ms expressing cells. Four proteins were in the "stress response" category: Lon protease (MSMEG_3582), an ATP-

dependent RNA helicase DEAD/DEAH box family protein (MSMEG_5042), sigma factor SigH (MSMEG_1914) and alkylhydroperoxide reductase (MSMEG_4891). Lon is a stress-responsive protease. Since it is known to degrade all TA system antitoxins in *Escherichia coli* (Gerdes and Maisonneuve, 2012), it may also have an analogous role and cleave the MazE-ms antitoxin in *M. smegmatis*. Uniprot places the MSMEG_5042 RNA helicase as functioning in ribosome biogenesis, mRNA degradation and translation initiation. The alternate sigma factor SigH is activated by oxidative, heat and nitric oxide stress (Sharp et al., 2016) while Uniprot places alkylhydroperoxide reductase MSMEG_4891 in protection from oxidative damage by detoxifying peroxides.

Consistent with the sustained protein synthesis while MazF-ms was being expressed, we observed new synthesis of a subset of elongation factors and ribosomal proteins. Four elongation factors were upregulated: EF-Ts (MSMEG_2520), EF-Tu (MSMEG_1401), EF-P (MSMEG_3035) and BipA/TypA ribosome-binding GTPase (MSMEG_5132). Many of these elongation factors are associated with the bacterial stress response. Finally, 18 of the 51 ribosomal proteins were also upregulated. The other 33 ribosomal protein levels were stable.

DISCUSSION

Bacterial genomes are under constant pressure to remain compact while also retaining genes that provide a competitive edge for survival in their natural environments. Acquisition of TA systems in bacterial genomes is thought to represent one potent vehicle for stress protection. In contrast to the ~90 TA systems in its pathogenic relative *M. tuberculosis*, the *M. smegmatis* genome harbors just three TA systems (MazEF, PhD-Doc, and VapBC) as one facet of its stress survival armamentarium. A thorough understanding on how a TA system acts to protect its host from stress requires determination of the function of the toxin, and the function of the toxin is informed by determining its intracellular target. Since all MazF toxins are generally single-strand and sequence-specific endoribonucleases, in this work we identified the RNA target of the MazF-ms toxin using 5' RNA-seq which revealed its detailed mechanism of action. To our surprise, MazF-ms did not behave like the vast majority of MazF toxins that appear to predominantly cleave mRNAs (reviewed in Masuda and Inouye, 2017). Instead, MazF-ms behaves almost exactly like the only other known exception, MazF-mt9, one of the 11 MazF family members in *M. tuberculosis* (Barth et al., 2019).

MazF-mt9 is an outlier because it requires both structure and sequence for its highly specific recognition of a single tRNA isoacceptor. This requirement for structure is much like VapC toxins, even though MazF-mt9 and VapC toxins lack sequence or structural similarity (Cruz et al., 2015; Schifano et al., 2016; Winther et al., 2016; Schifano and Woychik, 2017; Cintrón et al., 2019). MazF-ms now represents the second example of a MazF toxin that targets tRNA for cleavage, thus reducing the levels of only this tRNA species *in vivo*. Since we were able to unequivocally detect MazF-ms target RNAs with the required 5'-OH using 5' RNA-seq, there was no apparent masking of the precise cleavage site due to the presence of an RNA modification. This surgical depletion of just one tRNA results in ribosome stalling at codons requiring this depleted tRNA and proteome remodeling to favor sustained synthesis of only AAA-deficient proteins (**Figure 8**).

Why is Lys tRNA such an important conduit for proteome remodeling, ostensibly during stress, in mycobacteria? It appears that many proteins critical for stress survival in these GC rich genomes (67.4% GC in *M. smegmatis;* 65.6% in *M. tuberculosis*) are deficient in AAA Lys codons, allowing for their sustained translation while the cell can save energy by not synthesizing proteins that do not contribute to this endpoint. Indeed, as discussed above, many of the upregulated proteins annotated in the stress response category have direct roles in one or more stress responses: cold shock, heat shock, oxidative stress and nitric oxide stress. The stress-specific sigma factor SigH was also

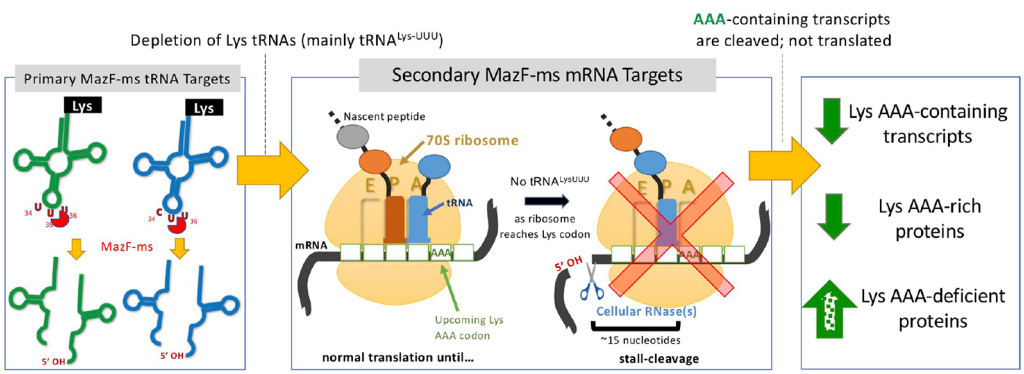

FIGURE 8 | Summary of the proposed MazF-ms mechanism of action. MazF-ms primarily targets tRNA$^{Lys-UUU}$ and tRNA$^{Lys-CUU}$, depleting their intracellular levels. Inactivation of these tRNAs lead to selective ribosome stalling predominantly in rare Lys-AAA codons, culminating in cleavage of the transcript. This results in codon-biased proteomic changes, favoring Lys-AAA-depleted transcripts.

upregulated. Therefore, it is implicated as the primary RNA polymerase sigma factor enlisted for the sustained transcription of *M. smegmatis* genes whose proteins were upregulated after toxin expression. However, other upregulated proteins in our dataset that are not commonly associated with the stress response, i.e. at least some elongation factors, appear to have indirect roles (reviewed in Starosta et al., 2014). EF-Tu is maximally expressed during stress in *E. coli* (Muela et al., 2008). Since the guanine nucleotide exchange factor EF-Ts assembles with EF-Tu in a 2:2 stoichiometry (Kawashima et al., 1996), it is expected to be expressed at levels equivalent to EF-Tu. EF-P rescues ribosomes stalled at poly-proline stretches. However, EF-P is only active when lysinylated or hydrozylysinylated (Doerfel et al., 2013; Ude et al., 2013). Thus, the lysine from cleaved Lys tRNAs might be recycled and used to activate EF-P. Finally, BipA/TypA, while not essential, appears to confer a growth advantage by regulating the synthesis of a subclass of proteins in cells under cold shock, low pH, oxidative stress, antimicrobial peptide stress and detergent stress.

DATA AVAILABILITY STATEMENT

The proteome data and RNA-seq data have been deposited and made public in the Massive database under the accession number MSV000084300 (https://massive.ucsd.edu/ProteoSAFe/dataset.jsp?task=1c14b9a1fd354c8fbb7ddcf03ace9fdd) and SRA under the accession number PRJNA564437 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA564437), respectively.

AUTHOR CONTRIBUTIONS

VB performed all experiments. Data analysis/interpretation was performed by VB. NW directed the project and acquired funding. NW and VB wrote the manuscript.

ACKNOWLEDGMENTS

We thank Haiyan Zheng and the Biological Mass Spectrometry Facility of Robert Wood Johnson Medical School and Rutgers (supported by National Center for Research Resources Grant S10OD016400) for quantitative mass spectrometry analysis and expertise, and the Genome Technology Center, New York University School of Medicine (partially supported by the Cancer Center Support Grant P30CA016087 at the Laura and Isaac Perlmutter Cancer Center) for providing Illumina sequencing services. This work was funded in part by a fellowship to VB from the Coordinating Agency for Advanced Training of Graduate Personnel (CAPES) within the Ministry of Education of Brazil, a Louis Bevier Fellowship from Rutgers University to VB, and National Institutes of Health grant R21 AI123859 to NW. VB is also a recipient of the Aaron Shatkin Scholarship from Rutgers University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.01356/full#supplementary-material

REFERENCES

Barth, V. C., Zeng, J.-M., Vvedenskaya, I. O., Ouyang, M., Husson, R. N., and Woychik, N. A. (2019). Toxin-mediated ribosome stalling reprograms the mycobacterium tuberculosis proteome. *Nat. Commun.* 10, 3035. doi: 10.1038/s41467-019-10869-8

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Cintrón, M., Zeng, J.-M., Barth, V. C. V. C., Cruz, J. W. J. W., Husson, R. N. R. N., and Woychik, N. A. N. A. (2019). Accurate target identification for mycobacterium tuberculosis endoribonuclease toxins requires expression in their native host. *Sci. Rep.* 9, 5949. doi: 10.1038/s41598-019-41548-9

Cruz, J. W., Sharp, J. D., Hoffer, E. D., Maehigashi, T., Vvedenskaya, I. O., Konkimalla, A., et al. (2015). Growth-regulating mycobacterium tuberculosis VapC-mt4 toxin is an isoacceptor-specific tRNase. *Nat. Commun.* 6, 1–12. doi: 10.1038/ncomms8480

Doerfel, L. K., Wohlgemuth, I., Kothe, C., Peske, F., Urlaub, H., and Rodnina, M. V. (2013). EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. *Science* 339, 85–88. doi: 10.1126/science.1229017

Ehrt, S., Guo, X. V., Hickey, C. M., Ryou, M., Monteleone, M., Riley, L. W., et al. (2005). Controlling gene expression in mycobacteria with anhydrotetracycline and tet repressor. *Nucleic Acids Res.* 33, e21. doi: 10.1093/nar/gni013

Frampton, R., Aggio, R. B. M., Villas-Bôas, S. G., Arcus, V. L., and Cook, G. M. (2012). Toxin-antitoxin systems of mycobacterium smegmatis are essential for cell survival. *J. Biol. Chem.* 287, 5340–5356. doi: 10.1074/jbc.M111.286856

Gerdes, K., and Maisonneuve, E. (2012). Bacterial persistence and toxin-antitoxin loci. *Annu. Rev. Microbiol.* 66, 103–123. doi: 10.1146/annurev-micro-092611-150159

Harms, A., Brodersen, D. E., Mitarai, N., and Gerdes, K. (2018). Toxins, targets, and triggers: an overview of toxin-antitoxin biology. *Mol. Cell* 70, 768–784. doi: 10.1016/j.molcel.2018.01.003

Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 37, 1–13. doi: 10.1093/nar/gkn923

Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57. doi: 10.1038/nprot.2008.211

Kawashima, T., Berthet-Colominas, C., Wulff, M., Cusack, S., and Leberman, R. (1996). The structure of the escherichia coli EF-Tu.EF-Ts complex at 2.5 A resolution. *Nature* 379, 511–518. doi: 10.1038/379511a0

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10, R25. doi: 10.1186/gb-2009-10-3-r25

Lowe, T. M., and Chan, P. P. (2016). tRNAscan-SE on-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res.* 44, W54–W57. doi: 10.1093/nar/gkw413

Masuda, H., and Inouye, M. (2017). Toxins of prokaryotic toxin-antitoxin systems with sequence-specific endoribonuclease activity. *Toxins (Basel)* 9, 140. doi: 10.3390/toxins9040140

Muela, A., Seco, C., Camafeita, E., Arana, I., Orruño, M., López, J. A., et al. (2008). Changes in escherichia coli outer membrane subproteome under environmental conditions inducing the viable but nonculturable state. *FEMS Microbiol. Ecol.* 64, 28–36. doi: 10.1111/j.1574-6941.2008.00453.x

Ramage, H. R., Connolly, L. E., and Cox, J. S. (2009). Comprehensive functional analysis of mycobacterium tuberculosis toxin-antitoxin systems: implications for pathogenesis, stress responses, and evolution. *PloS Genet.* 5, e1000767. doi: 10.1371/journal.pgen.1000767

Sala, A., Bordes, P., and Genevaux, P. (2014). Multiple toxin-antitoxin systems in mycobacterium tuberculosis. *Toxins (Basel).* 6, 1002–1020. doi: 10.3390/toxins6031002

Schifano, J. M., and Woychik, N. A. (2017). Cloaked dagger: tRNA slicing by an unlikely culprit. *RNA Biol.* 14, 15–19. doi: 10.1080/15476286.2016.1255396

Schifano, J. M., Vvedenskaya, I. O., Knoblauch, J. G., Ouyang, M., Nickels, B. E., and Woychik, N. A. (2014). An RNA-seq method for defining endoribonuclease cleavage specificity identifies dual rRNA substrates for toxin MazF-mt3. *Nat. Commun.* 5, 3538. doi: 10.1038/ncomms4538

Schifano, J. M., Cruz, J. W., Vvedenskaya, I. O., Edifor, R., Ouyang, M., Husson, R. N., et al. (2016). tRNA is a new target for cleavage by a MazF toxin. *Nucleic Acids Res.* 44, 1256–1270. doi: 10.1093/nar/gkv1370

Sharp, J. D., Singh, A. K., Park, S. T., Lyubetskaya, A., Peterson, M. W., Gomes, A. L. C., et al. (2016). Comprehensive definition of the SigH regulon of mycobacterium tuberculosis reveals transcriptional control of diverse stress responses. *PloS One* 11, e0152145. doi: 10.1371/journal.pone.0152145

Song, S., and Wood, T. K. (2018). Post-segregational killing and phage inhibition are not mediated by cell death through toxin/antitoxin systems. *Front. Microbiol* 9, 814. doi: 10.3389/fmicb.2018.00814

Starosta, A. L., Lassak, J., Jung, K., and Wilson, D. N. (2014). The bacterial translation stress response. *FEMS Microbiol. Rev.* 38, 1172–1201. doi: 10.1111/1574-6976.12083

Strimmer, K. (2008). fdrtool: a versatile r package for estimating local and tail area-based false discovery rates. *Bioinformatics* 24, 1461–1462. doi: 10.1093/bioinformatics/btn209

Ude, S., Lassak, J., Starosta, A. L., Kraxenberger, T., Wilson, D. N., and Jung, K. (2013). Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. *Science* 339, 82–85. doi: 10.1126/science.1228985

Winther, K., Tree, J. J., Tollervey, D., and Gerdes, K. (2016). VapCs of Mycobacterium tuberculosis cleave RNAs essential for translation. *Nucleic Acids Res.* 44, 9860–9871. doi: 10.1093/nar/gkw781

Wu, X., and Bartel, D. P. (2017). KpLogo: positional k -mer analysis reveals hidden specificity in biological sequences. *Nucleic Acids Res.* 45, W534–W538. doi: 10.1093/nar/gkx323

Xie, Y., Wei, Y., Shen, Y., Li, X., Zhou, H., Tai, C., et al. (2018). TADB 2.0: an updated database of bacterial type II toxin–antitoxin loci. *Nucleic Acids Res.* 46, D749–D753. doi: 10.1093/nar/gkx1033

Yamaguchi, Y., Park, J.-H., and Inouye, M. (2011). Toxin-antitoxin systems in bacteria and archaea. *Annu. Rev. Genet.* 45, 61–79. doi: 10.1146/annurev-genet-110410-132412

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

ORIGINAL RESEARCH
published: 27 February 2020
doi: 10.3389/fgene.2020.00097

Comprehensive Functional Analysis of *Escherichia coli* Ribosomal RNA Methyltransferases

Philipp Pletnev[1,2,3†], *Ekaterina Guseva*[4†], *Anna Zanina*[1†], *Sergey Evfratov*[1], *Margarita Dzama*[4], *Vsevolod Treshin*[1], *Alexandra Pogorel'skaya*[4], *Ilya Osterman*[1,2], *Anna Golovina*[5], *Maria Rubtsova*[1,2], *Marina Serebryakova*[2,5], *Olga V. Pobeguts*[6], *Vadim M. Govorun*[6], *Alexey A. Bogdanov*[1,5], *Olga A. Dontsova*[1,2,3,5] *and Petr V. Sergiev*[1,2,5,7*]

[1] Department of Chemistry, Lomonosov Moscow State University, Moscow, Russia, [2] Center of Life Sciences, Skolkovo Institute of Science and Technology, Moscow, Russia, [3] Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Moscow, Russia, [4] Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University, Moscow, Russia, [5] Belozersky Institute of Physico-Chemical Biololgy, Lomonosov Moscow State University, Moscow, Russia, [6] Federal Research and Clinical Centre of Physical-Chemical Medicine, Moscow, Russia, [7] Institute of Functional Genomics, Lomonosov Moscow State University, Moscow, Russia

OPEN ACCESS

Edited by:
Michael Ibba,
The Ohio State University,
United States

Reviewed by:
Hyouta Himeno,
Hirosaki University, Japan
Dominique Belin,
Université de Genève, Switzerland

***Correspondence:**
Petr V. Sergiev
petya@genebee.msu.ru

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: 14 May 2019
Accepted: 29 January 2020
Published: 27 February 2020

Citation:
Pletnev P, Guseva E, Zanina A, Evfratov S, Dzama M, Treshin V, Pogorel'skaya A, Osterman I, Golovina A, Rubtsova M, Serebryakova M, Pobeguts OV, Govorun VM, Bogdanov AA, Dontsova OA and Sergiev PV (2020) Comprehensive Functional Analysis of Escherichia coli Ribosomal RNA Methyltransferases. Front. Genet. 11:97. doi: 10.3389/fgene.2020.00097

Ribosomal RNAs in all organisms are methylated. The functional role of the majority of modified nucleotides is unknown. We systematically questioned the influence of rRNA methylation in *Escherichia coli* on a number of characteristics of bacterial cells with the help of a set of rRNA methyltransferase (MT) gene knockout strains from the Keio collection. Analysis of ribosomal subunits sedimentation profiles of the knockout strains revealed a surprisingly small number of rRNA MT that significantly affected ribosome assembly. Accumulation of the assembly intermediates was observed only for the *rlmE* knockout strain whose growth was retarded most significantly among other rRNA MT knockout strains. Accumulation of the 17S rRNA precursor was observed for *rsmA* (*ksgA*) knockout cells as well as for cells devoid of functional *rsmB* and *rlmC* genes. Significant differences were found among the WT and the majority of rRNA MT knockout strains in their ability to sustain exogenous protein overexpression. While the majority of the rRNA MT knockout strains supported suboptimal reporter gene expression, the strain devoid of the *rsmF* gene demonstrated a moderate increase in the yield of ectopic gene expression. Comparative 2D protein gel analysis of rRNA MT knockout strains revealed only minor perturbations of the proteome.

Keywords: ribosome, rRNA, methyltransferase, modification, *Escherichia coli*

INTRODUCTION

Ribosomal RNA (rRNA) of all organisms contains modified nucleosides (Sergiev et al., 2011). The functional role of such modifications is unknown in many cases. Some modifications were shown to be required for proper ribosome assembly (Connolly et al., 2008), interaction with ribosomal ligands (Burakovsky et al., 2012), antibiotic resistance (Weisblum, 1995), and the correct functioning of

particular regulatory mechanisms (Vazquez-Laslop et al., 2008; Vazquez-Laslop et al., 2010; Prokhorova et al., 2013). In 2012, the entire list of *E. coli* rRNA methyltransferase genes was completed (Golovina et al., 2012). Not a single modification of rRNA was found to be essential for bacterial cell survival, although earlier studies indicated that the lack of several modifications might cumulatively have a deleterious effect on ribosome activity (Green and Noller, 1996). Obviously, none of the rRNA modifications is required for the general ability of a ribosome to synthesize proteins. However, in a living cell, a ribosome should synthesize the proteins quickly, at the right proportions, and in a cost-effective manner (Li et al., 2014). Ribosomal RNA modification might contribute to the fine-tuning of particular gene expression mechanisms (Sergiev et al., 2012; Prokhorova et al., 2013) or contribute to the efficiency of protein biosynthesis in general. At favorable growth conditions, the inefficiency of protein biosynthesis might be tolerated, while an increased load on the protein biosynthesis machinery, such as in the artificial case of foreign gene overexpression, or when it is necessary to replace damaged proteins, modification of rRNA might play a role. In this work we systematically studied the influence of rRNA methyltransferase gene knockouts (**Table 1**) on bacterial growth, the accumulation of assembly intermediates, deviations in the proteome, and the ability to sustain excessive protein synthesis.

RESULTS

Expression of rRNA Methyltransferase Genes at Different Stages of Bacterial Culture Growth

Gene expression analysis in specific conditions might give a hint about its specialized function. Various genes, coding for rRNA modification enzymes, show distinct expression patterns according to the databases on a global analysis of gene

TABLE 1 | List of *E. coli* rRNA MT coding genes and the phenotypes of their knockouts.

Nucleotide	Enzyme	Reference	Growth of knockout strain/ growth at overexpression*	CER expression**	FastFT expression***	Accumulation of assembly intermediates and 17S rRNA precursor	Expression level**** and timing
16S rRNA							
527 m^7G	RsmG	Okamoto et al., 2007	+++\++	+++	+++	–/–	+++, 3h
966 m^2G	RsmD	Lesnyak et al., 2007	+++\+++	+	+++	–/–	+++, 2h
967 m^5C	RsmB	Tscherne et al., 1999a; Gu et al., 1999	+++\+++	++	++	–/++	++++, 4h
1207 m^2G	RsmC	Tscherne et al., 1999b	+++\+++	++	+++	–/–	++++, 2h
1402 m^4Cm	RsmI, RsmH	Kimura and Suzuki, 2010	+++\+	+	+/+++#	–/–	+, 2h
1407 m^5C	RsmF	Andersen and Douthwaite, 2006	+++\+++	++++	++++	–/–	+++, 2h
1498 m^3U	RsmE	Basturea et al., 2006	+++\+++	+++	+++	–/–	+++, 2h
1516 m^2G	RsmJ	Basturea et al., 2012	+++\+++	+	++	–/–	+++, 2h
1518/9 $m^6{}_2A$	RsmA	Helser et al., 1972; Poldermans et al., 1979	+++\+++	+	+	–/++	++++, 2h
23S rRNA							
745 m^1G	RlmA	Gustafsson and Persson, 1998	+++\++	+	+/+++#	–/–.	++++, 2h
747 m^5U	RlmC	Madsen et al., 2003	+++\++	+	+++	–/++	++, 2h
1618 m^6A	RlmF	Sergiev et al., 2008	+++\++	+	+++	–/–	+++, 1–7h
1835 m^2G	RlmG	Sergiev et al., 2006	+++\+++	++	++	–/–	+++, 2h
1915 $m^3\Psi$	RlmH	Purta et al., 2008a; Ero et al., 2008	+++\+++	+	++	–/–	+++, 2h
1939 m^5U	RlmD	Agarwalla et al., 2002; Madsen et al., 2003	+++\+++	+++	+++	–/–	+++, 2h
1962 m^5C	RlmI	Purta et al., 2008b	+++\+	+	++	–/–	+++, 2h
2030 m^6A	RlmJ	Golovina et al., 2012	+++\+++	+	+++	–/–	++++, 2h
2069 m^7G, 2445 m^2G	RlmKL	Kimura et al., 2012	+++\+++	+	+++	–/–	+++, 2h
2251 Gm	RlmB	Lovgren and Wikstrom, 2001	+++\+++	++	+++	–/–	+++, 2h
2498 Cm	RlmM	Purta et al., 2009	+++\+++	+	++	+?/–	+++, 2h
2503 m^2A	RlmN	Toh and Mankin, 2008	+++\++	+	+++	+?/-	+++, 2h
2552 Um	RlmE	Caldas et al., 2000; Bugl et al., 2000	+\+	+	+/++#	+++/–	++++, 3h
	WT		+++	+++	+++	–/–	

*+++ corresponds to doubling times 40–60 min, ++ 60–70 min, + > 70 min. ** CER fluorescence after overnight growth at inducing conditions ++++ > $6{\cdot}10^5$, +++ $4–6{\cdot}10^5$, ++ $2–4{\cdot}10^5$, + < $4{\cdot}10^5$. ***Fluorescence intensity of the blue, newly produced form of the FastFT at the exponential growth conditions ++++ > $2{\cdot}10^3$, +++ $1–2{\cdot}10^3$, ++ $5{\cdot}10^2–10^3$, + < $5{\cdot}10^2$. ****Expression levels at a timepoint with maximal expression. ++++ > 10^{-4}, +++ > 10^{-5}, ++ > 10^{-6}, + < 10^{-6} relative to the 16S rRNA. #During growth of these strains transformed by FastFT expression plasmid a significant amount of non-fluorescent cells were accumulated.*

expression in a variety of conditions (Sergiev et al., 2012). We monitored expression of rRNA methyltransferase genes experimentally at different stages of bacterial culture growth using RT-qPCR (**Figure 1**). In the computer analysis of GEO profiles of gene expression focused on rRNA modification genes (Sergiev et al., 2012) we noticed that the majority of rRNA methylation genes, except for *rsmA*(*ksgA*) and *rlmE*, are coexpressed with other genes associated with fast growth. Here, we demonstrated experimentally that the expression of the majority of rRNA methyltransferase genes is induced at an early exponential phase (**Figure 1A**). The amount of mRNA molecules coding for different rRNA methyltransferases

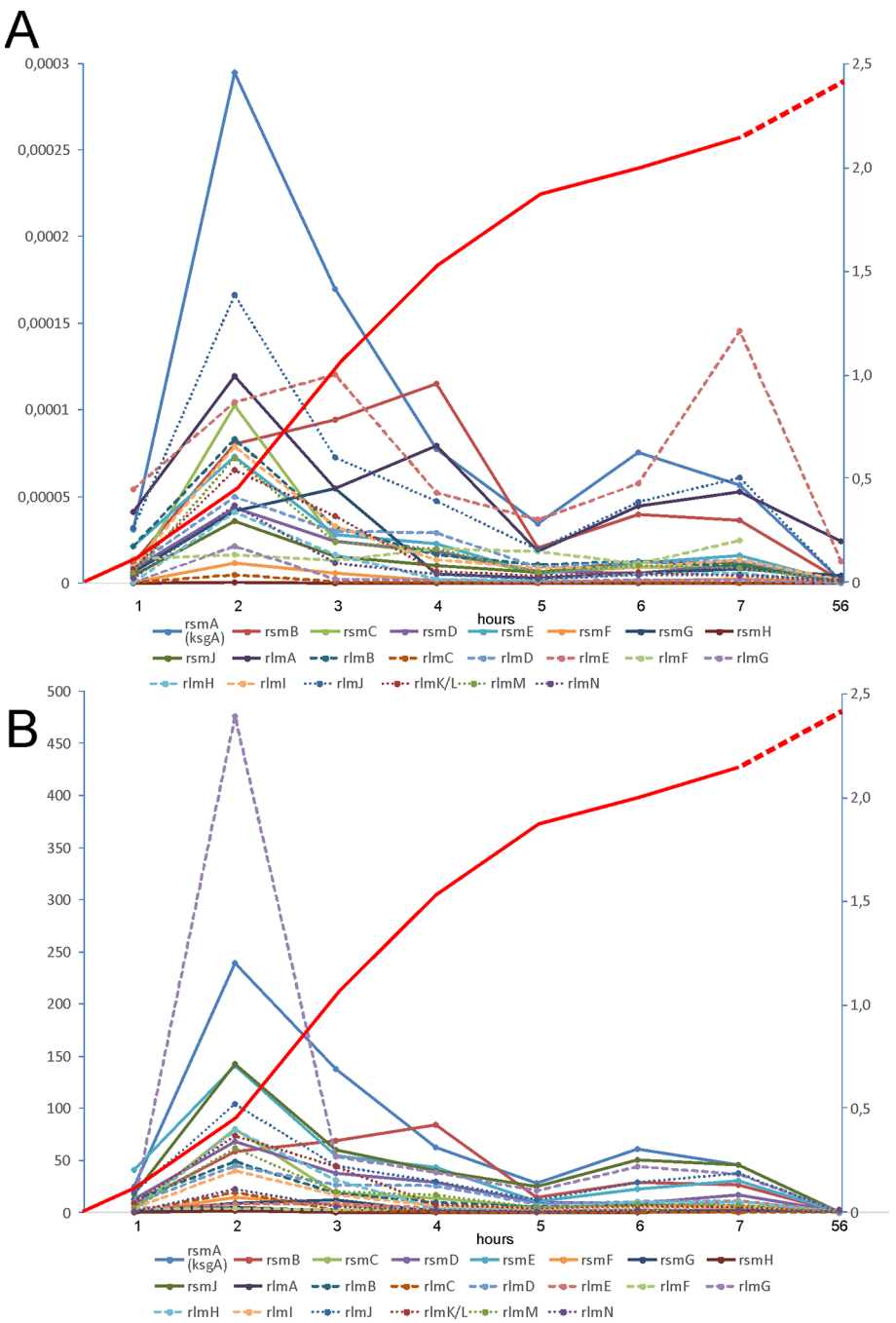

FIGURE 1 | Expression of rRNA methyltransferase genes at different phases of *E. coli* culture growth. **(A)** Amounts of rRNA methyltransferases mRNA relative to the amount of 16S rRNA as revealed by RT qPCR (left scale). **(B)** Change in the proportion of rRNA methyltransferases mRNA to the 16S rRNA relative to that in the stationary phase (left scale). The keys to the graphs are shown below the panels. Red curves (right scale) correspond to the A_{600} of the cell culture.

normalized to the amount of rRNA differs by three orders of magnitude. We observed the highest expression level for the *rsmA*(*ksgA*) gene and the lowest for *rsmH*. Analysis of the growth phase dependencies of expression (**Figure 1B**) revealed a group of rRNA methyltransferase genes, namely, *rlmN*, *rlmC*, *rsmF*, *rsmG*, *rlmA*, *rlmF*, and *rsmH*, with nearly constitutive expression. The remaining rRNA MT coding genes demonstrated a variable extent of induction from small to moderate for *rsmC*, *rlmH*, *rlmKL*, *rsmD*, *rlmM*, *rlmB*, *rlmD*, *rlmI*, high for *rsmE*, *rsmJ*, *rlmJ*, and extreme for *rlmG* and *rsmA*(*ksgA*). Genes coding for *rsmB* and *rlmE* have specific expression patterns (**Figure 1**). The *RsmB* gene is transcribed at a later growth stage, the maximal expression level being attained at 4 h after dilution of the culture in fresh media. The gene coding for *RlmE* is transcribed almost constitutively with the maximal level of expression reached after 7 h of bacterial culture dilution in fresh media.

Influence of rRNA Methyltransferase Genes Inactivation on the Accumulation of Ribosome Assembly Intermediates

A function of rRNA modification enzymes in ribosome assembly was proposed for RlmE and RsmA(KsgA) rRNA methyltransferases. The knockout of the former caused an accumulation of assembly intermediates and slow growth (Bugl et al., 2000; Caldas et al., 2000; Hager et al., 2002; Arai et al., 2015), which could be suppressed by overexpression of small GTPases Obg and EngA (Tan et al., 2002). RsmA(KsgA) was proposed to be the ribosome biogenesis factor, utilizing its methyltransferase activity to trigger its dissociation upon successful completion of the 30S subunit assembly (Connolly et al., 2008). Involvement of other rRNA methyltransferases in the ribosome assembly pathway could be hypothesized (Sergiev et al., 2011). We used a collection of rRNA methyltransferase knockout strains (Baba et al., 2006) to systematically study the accumulation of assembly intermediates of ribosomal subunits. Knockout strains were grown in a rich medium at 37°C and 20°C. Low temperature slows conformational rearrangements of RNA and is known to aggravate ribosome assembly defects (Shajani et al., 2011); thus, at 20°C we expected to reveal more severe defects then at 37°C. Sucrose density centrifugations of the rRNA methyltransferase gene knockout cell lysates were done at low, 1 mM (**Figure 2A**) and high 10 mM (**Figure 2B**) magnesium ion concentrations to reveal any differences in accumulation of assembly intermediates or subunit association. Only the *rlmE* knockout strain previously known to accumulate a significant amount of assembly intermediates (Bugl et al., 2000; Caldas et al., 2000; Hager et al., 2002; Arai et al., 2015) demonstrated such an effect at both magnesium ion concentrations. In addition, small peaks of presumably assembly intermediates might be observed on the

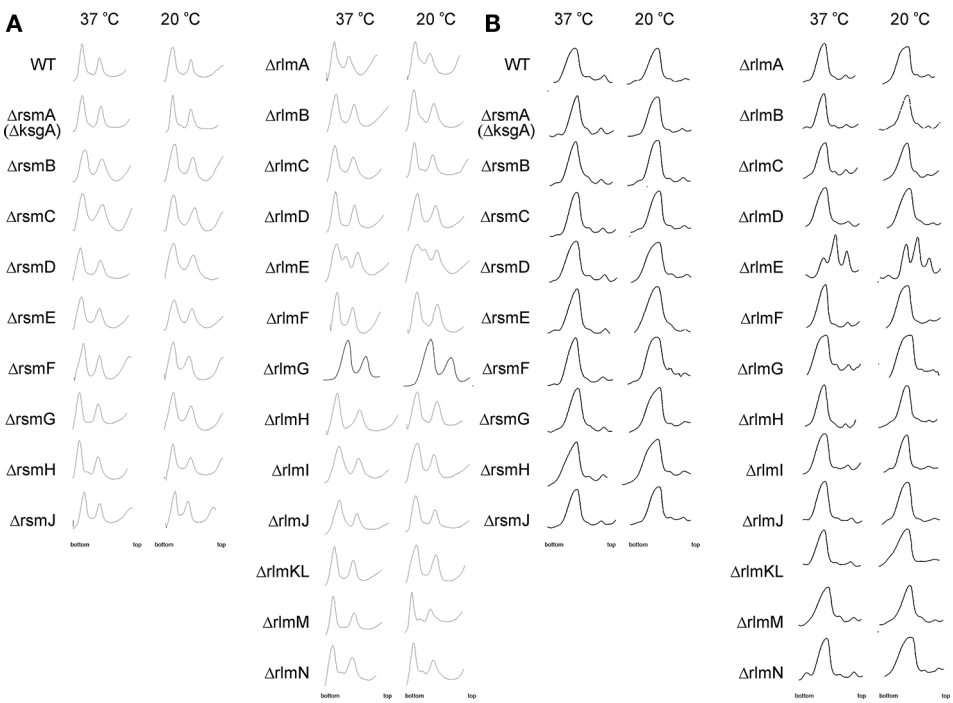

FIGURE 2 | Accumulation of ribosomal subunits assembly intermediates in the *E. coli* strains with inactivation of the 16S rRNA and 23S rRNA methyltransferase genes marked on the left side. WT corresponds to the parental isogeneic strain carrying all set of rRNA methyltransferase genes. Shown are sucrose gradient centrifugation profiles at subunit dissociation conditions (1 mM magnesium ions concentration) **(A)** and association conditions (10 mM magnesium ions concentration) **(B)**. Left panels correspond to the cells grown at 37°C, while that to the right correspond to the cells grown at 20°C known to exacerbate ribosomal subunits assembly defects.

sucrose gradients at a low magnesium ion concentration of the extracts from Δ*rlmM* and Δ*rlmN* strains.

Surprisingly, none of the other rRNA methyltransferase knockout strains demonstrated the accumulation of ribosomal subunits assembly intermediates with sedimentation properties different from that of mature ribosomal subunits, arguing against the essentiality of the corresponding rRNA modifications for ribosome assembly.

Nucleolytic processing of the 17S rRNA precursor resulting in the formation of the 16S rRNA happens in the late stage of small ribosomal subunit assembly (Smith et al., 2018). For the few strains deficient in rRNA modification enzymes (Gutgsell et al., 2005; Connolly et al., 2008) excessive accumulation of the 17S rRNA precursor was demonstrated. However, up to now, no systematic study of rRNA modification's influence on the processing of the 17S rRNA precursor was available. We decided to analyze 17S rRNA to 16S rRNA ratio in the cells devoid of each of the rRNA methyltransferases (**Figure 3**) with the help of the RT qPCR. An accumulation of the 17S rRNA precursor was detected for cells of the Δ*rsmA*(Δ*ksgA*) strain grown at a low temperature, in agreement with previously published data (Connolly et al., 2008). In addition, a substantial increase in the amount of the 17S rRNA processing intermediate was detected for the cells of the Δ*rsmB* and Δ*rlmC*

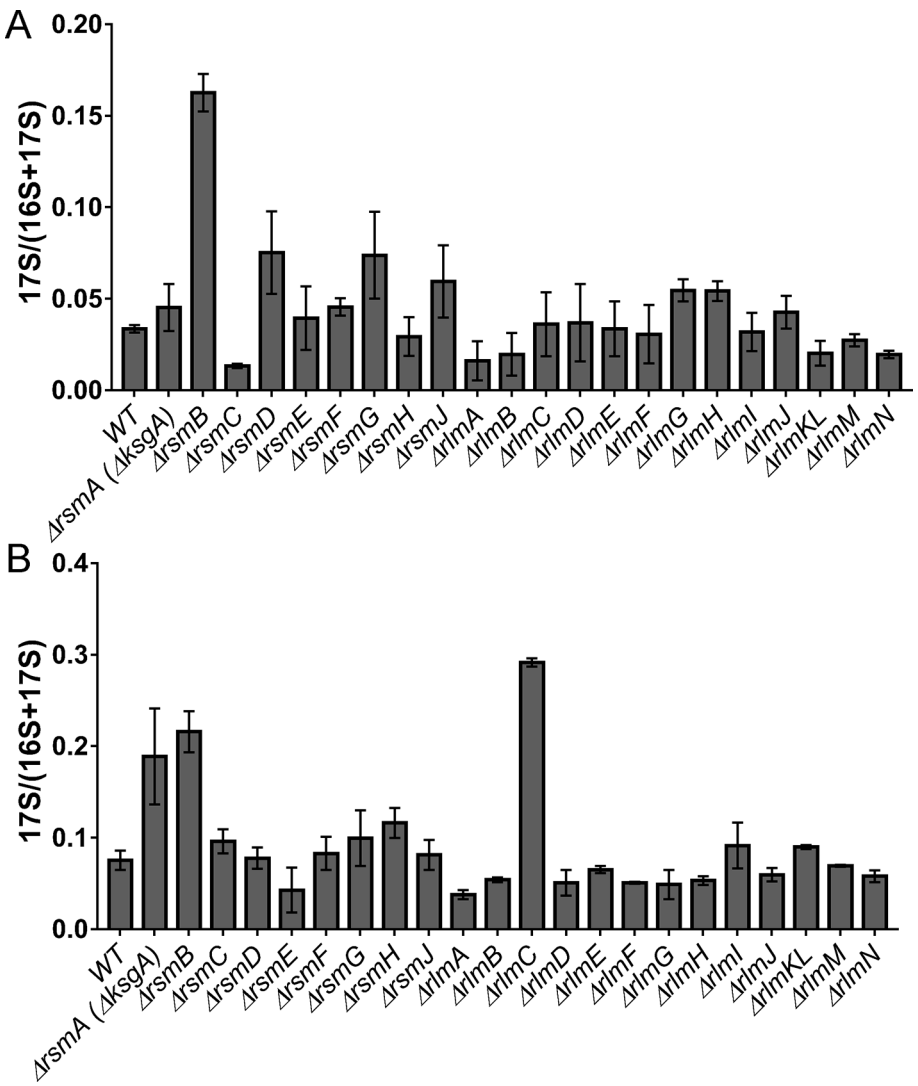

FIGURE 3 | Accumulation of 17S rRNA precursor in the *E. coli* strains with inactivation of the 16S rRNA and 23S rRNA methyltransferase genes. Shown is the 17S rRNA to 16S+17S rRNA ratio determined by RT qPCR. Knockout strains are marked below the bars. (Panel **A**) corresponds to the cells grown at the 37°C, while (panel **B**) to the cells grown at the 20°C.

strains. For the former, a higher concentration of the precursor was detected at both tested temperatures, while for the latter, 17S was accumulated at 20°C.

Influence of rRNA Methyltransferase Gene Inactivation on the Growth Rate of Bacteria

Protein biosynthesis requires a large share of cellular resources. Any deviation from optimal protein synthesis efficiency should influence the doubling time of bacteria. We consistently measured the growth rates of bacteria lacking rRNA methyltransferase genes (**Figure 4**). At optimal conditions in a rich medium (**Figure 4A**) only the Δ*rlmE* strain demonstrated a significant, nearly twofold increase in the doubling time. Such a result correlates well with the accumulation of assembly intermediates in this strain.

Overexpression of an exogenous protein leads to the unproductive waste of biosynthetic protein resources. It may lead to significant growth retardation if the protein biosynthesis machinery would operate at suboptimal efficiency. To evaluate the influence of exogenous gene expression on cell growth, we used a reporter plasmid carrying the red fluorescent protein (*rfp*) gene under a constitutive T5 promoter and the cerulean fluorescent protein (*cer*) gene under the inducible Tet promoter. Growth rates were measured for rRNA methyltransferase knockout strains upon overexpression of the CER fluorescent protein gene (**Figure 4B**). More knockout strains revealed the difference from the wild-type parental strain in doubling time; in addition to the strain lacking

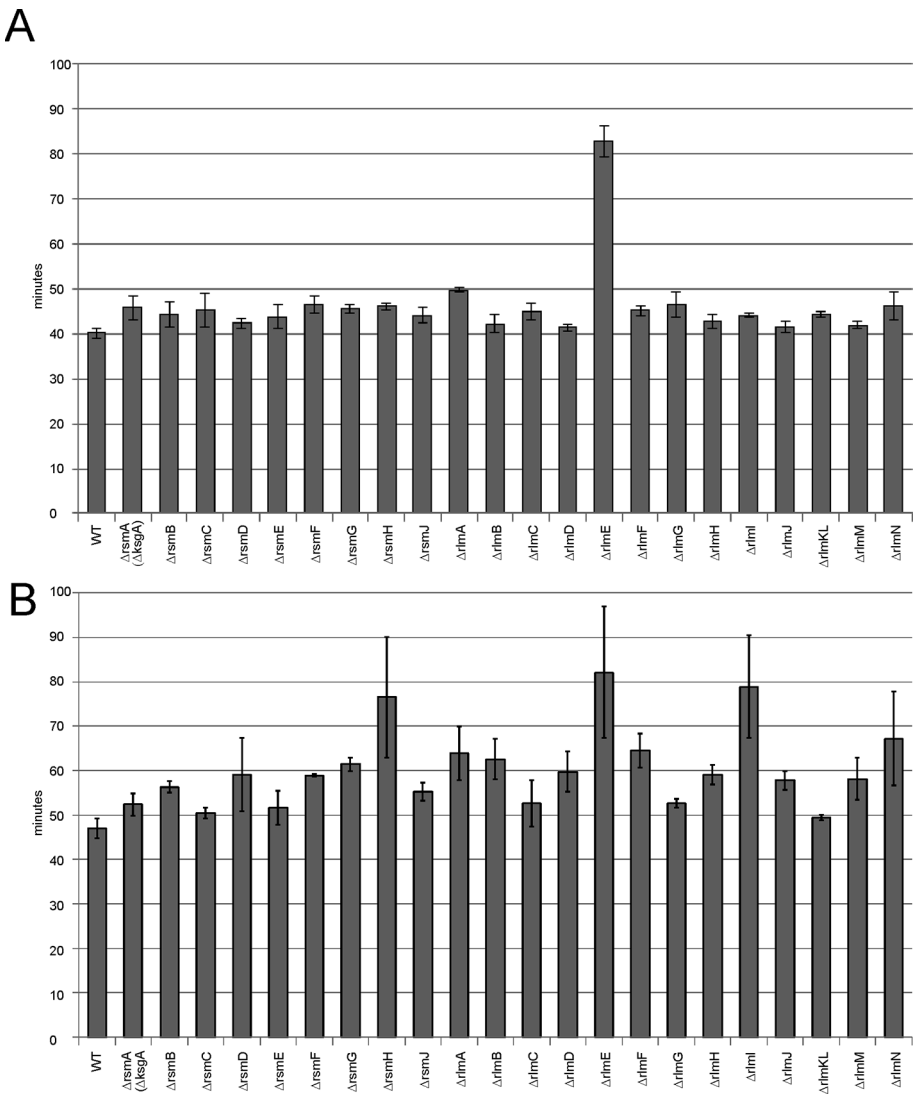

FIGURE 4 | Influence of rRNA methyltransferase gene inactivation on the growth rates of bacteria. **(A)** Doubling times of the strains devoid of the rRNA methyltransferase genes. Inactivated genes are labeled below the graph. WT corresponds to the parental isogeneic strain carrying a whole set of rRNA methyltransferase genes. **(B)** Doubling times of the same strains carrying a plasmid coding for the cerulean fluorescent protein (CER) and red fluorescent protein (RFP) reporter proteins at induction of reporter gene expression.

RlmE methyltransferase, significant growth retardation was found in the strains lacking RsmH and RlmI rRNA MTs.

Influence of rRNA Methyltransferase Gene Inactivation on the Efficiency of Constitutive and Induced Exogenous Protein Synthesis

In a living cell, different mRNA species compete for the protein biosynthesis machinery. If ribosomes and other components of the translation apparatus are present in excess over the total mRNA, then the transcriptional control of gene expression might function optimally and an increase in certain mRNA synthesis would not limit the expression of other genes. However, if the number of active ribosomes is low or they are functioning inefficiently, then excessive transcription of a gene would have a negative influence on other genes' expression due to the competition. To evaluate the influence of rRNA methylation on the availability of protein biosynthesis machinery for exogenous protein synthesis, we used a set of rRNA methyltransferase knockout strains transformed with a plasmid encoding for RFP under the constitutive T5 phage promoter and the *cer* gene under the inducible Tet promoter. The expression level of both genes could be monitored simultaneously in the rRNA MT knockout cells and compared to that of the wild-type cells. Thus, the overall ability of cells to support excessive protein synthesis might be deduced from this experiment as well as the influence of induced gene expression on the expression level of other genes, exemplified by a constitutively expressed *rfp* gene (**Figure 5**). The wild-type cells efficiently expressed both fluorescent proteins, and no reduction in constitutively transcribed *rfp* gene expression was observed upon additional expression of the *cer* gene. The yield of exogenous reporter protein synthesis was reduced for rRNA knockout strains. Only for the Δ*rsmB*, Δ*rsmC*, Δ*rsmG*, and Δ*rlmD* strains this reduction was

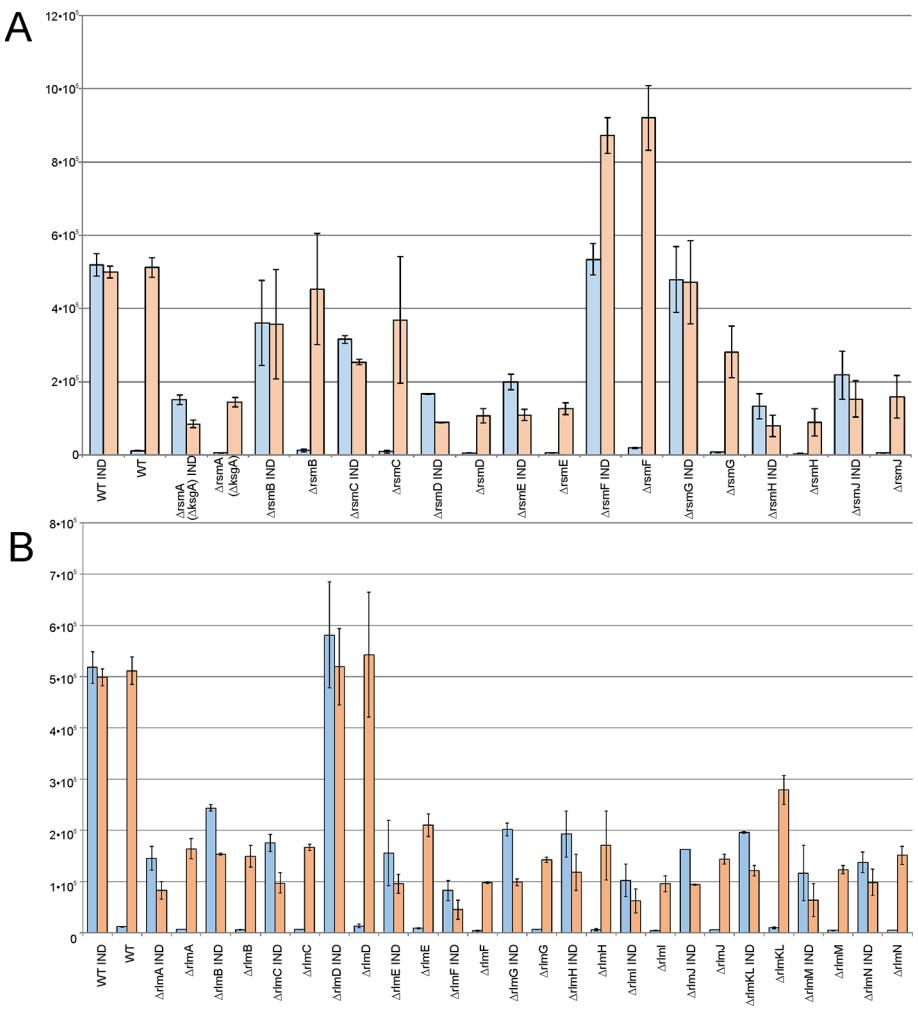

FIGURE 5 | Influence of rRNA methyltransferase gene inactivation on the expression of exogenous reporter genes encoded on the plasmid carrying red fluorescent protein (RFP) gene under the control of a constitutive T5 promoter and cerulean fluorescent protein (CER) gene under a control of an induced Tet promoter. Shown are the intensities of the CER (blue) and RFP (orange) fluorescence in the overnight cultures of the strains devoid of the 16S rRNA **(A)** and 23S rRNA **(B)** methyltransferase genes as labeled below the graphs. WT corresponds to the parental isogeneic strain carrying all set of rRNA methyltransferase genes. IND marks the graphs corresponding to the induction of CER gene expression by anhydrotetracycline.

mild, and not exceeding twofold. More significant, up to a 10-fold reduction of RFP expression was attributed to the lack of the remaining rRNA methyltransferases (**Figure 5**). Of note, a reduction in constitutively transcribed *rfp* gene expression was observed upon induction of *cer* gene transcription in the strains deficient in *rsmA*(*ksgA*), *rlmA, rlmC*, *rlmE*, *rlmF, rlmG*, *rlmJ*, *rlmKL*, *rlmM*, and *rlmN*. The only exception in this rule is the Δ*rsmF* strain supporting twofold higher expression of the RFP gene, relative to that of the wild-type strain, as well as mildly increased expression of the CER gene, relative to that in the wild-type strain at the uninduced state, and upon CER gene induction. To be sure that this observation was well reproduced, we repeated this experiment in 24 independent cell culture replicates (**Figure S1**). While the absolute level of CER and RFP fluorescence varied from culture to culture, the tendency always remained the same. In the strain devoid of RsmF, both CER and RFP levels were higher when those in the wild-type strain. To check whether this increase in expression is due to transcription or translation enhancement, we compared the amounts of *cer* and *rfp* mRNAs (**Figure S2**). An increase in the yield of both *cer* and *rfp* gene transcription in the Δ*rsmF* strain shows evidence that the overall upregulation of CER and RFP synthesis in this strain could be explained, at least partially, by the upregulation of transcription.

Our findings support a hypothesis that in the majority of strains lacking particular types of rRNA methylation, ribosome availability became a limiting factor for gene expression, where the strain devoid of RsmF methyltransferase appeared to be an exception.

Single Cell Analysis of the Efficiency of Protein Synthesis Upon the Inactivation of rRNA Methyltransferase Genes

To evaluate the capacity for exogenous protein synthesis in the individual cells of rRNA MT knockout strains, we applied a reporter plasmid encoding a FastFT fluorescent timer protein (Subach et al., 2009) and detected the intensity of fluorescence by the cell sorter. The FastFT protein undergoes a two-step maturation process in a way that a rapid formation of a blue fluorophore is followed by a slow conversion to a red fluorescent form with a half conversion rate of 7 h. With this rate, stationary phase cells, after 24 h of growth, contain mainly the fully converted red fluorescent form of FastFT, while rapidly growing cells contain the blue fluorescent form of the protein. We transformed rRNA MT knockout strains systematically with the plasmid coding for the FastFT protein under a control of an arabinose-inducible promoter and monitored the fluorescence at 405/460 nm for a newly synthesized blue fluorescent form and at 555/610 nm for a fully converted red fluorescent form (**Figure 6**). Monitoring of the FastFT protein level and synthesis was done throughout the growth curve of the bacterial culture.

At the stationary phase of the bacterial culture, the level of the predominant red form of the FastFT protein for the majority of the rRNA MT knockout strains exceeded that of the wild type (**Figures 6A, B**). We hypothesize that this might be due to the increase of the average cell size, rather than a change in FastFT protein biosynthesis. An exceptional presence of the blue FastFT form in the stationary phase cells was documented for the strain

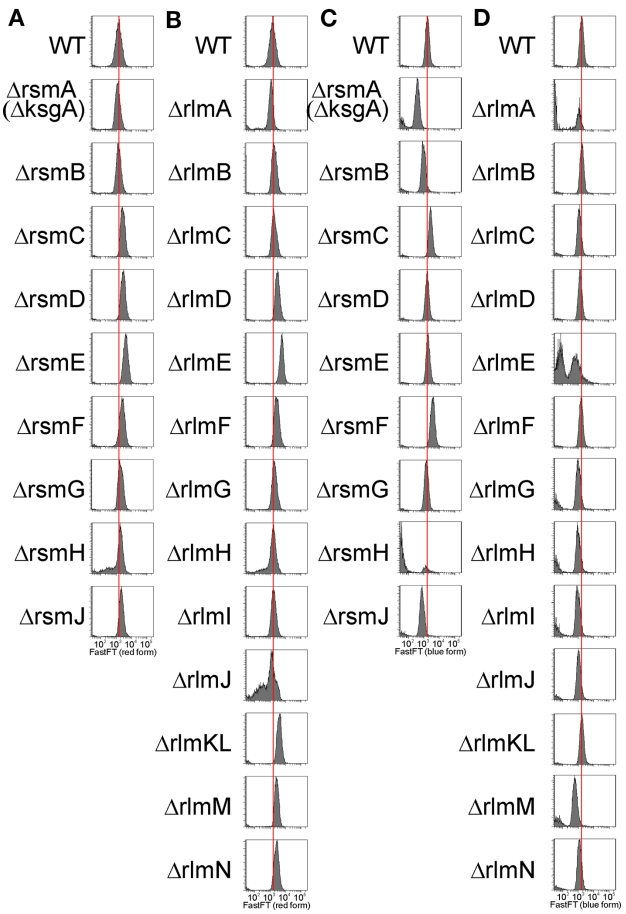

FIGURE 6 | Influence of rRNA methyltransferase gene inactivation on the expression of exogenous reporter gene FastFT. Shown is a distribution of cells (arbitrary units) by the red fluorescence intensity corresponding to the completely mature FastFT form in overnight cultures of the strains devoid of the 16S rRNA **(A)** and 23S rRNA **(B)** methyltransferase genes as labeled on the left side of the graphs. Distribution of cells (arbitrary units) by the blue fluorescence intensity corresponding to the newly synthesized FastFT form in the cultures at the exponential growth (5 h post dilution) of the strains devoid of the 16S rRNA **(C)** and 23S rRNA **(D)** methyltransferase genes as labeled on the left side of the graphs.

lacking the *rsmF* gene (**Figure S3**), which corresponds to a moderate increase in both CER and RFP reporter protein levels in this strain, relative to the wild-type strain described in the preceding section.

In the exponential growth phase (**Figures 6C, D**), the level of the predominant blue form of the FastFT protein was generally lower for the rRNA MT knockout strains, while the cell size was approximately the same. This is indicative of the decrease in exogenous protein synthesis efficiency upon rRNA MT gene inactivation, which is especially prominent in *rsmA*(*ksgA*), *rsmJ*, *rlmE*, and *rlmM* knockout cells. These findings correspond well with our data on the synthesis of CER and RFP reporter proteins discussed in the preceding section. Furthermore, a strain lacking the *rsmF* gene, and additionally the Δ*rsmC* strain, demonstrated

a higher amount of the freshly synthesized blue FastFT protein form in the exponentially growing cells, in line with an increased expression of other exogenous fluorescent protein genes.

Influence of rRNA Methyltransferase Gene Inactivation on the Composition of the Total Proteome

Lack of m^2G966/m^5C967 16S rRNA nucleotide modification perturbed the proteome of the cell and resulted in the misregulation of gene expression control by transcription attenuation (Prokhorova et al., 2013). Next, we sought to compare the proteomes of the strains lacking rRNA MT encoding genes. Proteins, whose abundance was changed in comparison with the parental strain, were detected by 2D protein gel electrophoresis (**Figure S4**). Surprisingly, none of the single rRNA methyltransferase gene knockouts led to a significant and reproducible perturbation of the proteome as revealed by 2D protein gel electrophoresis.

Three strains were selected for more detailed proteome analysis using a label-free shotgun proteome technique. The Δ*rlmE* strain demonstrated the most severe growth retardation, accumulation of ribosome assembly intermediates, and a significant reduction in the efficiency of reporter protein synthesis. The Δ*rsmF* strain was unique in its ability to support higher expression of the exogenous protein than the parental wild-type strain. The strain lacking the *rlmC* gene had a mild phenotypic abnormality and an increased level of the 17S rRNA precursor at low temperature. Very few proteins changed their abundance in the Δ*rlmC* strain (**Table 2**, **Table S1**). An analysis of transcription factors that might co-regulate genes whose expression depended on the RlmC revealed σS and ppGpp that regulate *sra* and *gadB* expression, although the statistical support is insufficient (Keseler et al., 2013). A more profound change was observed in the *rlmE* gene knockout. Visual inspection of the results suggests that σS, ppGpp, AppY, ArcA, and FNR activated genes are downregulated in the Δ*rlmE* strain, while genes repressed by PepA, ArcA, FNR, arginine, methionine, and pyrimidine nucleotides are upregulated. This statement, however, could get sufficient statistical support only for the transcription factor PepA (pvalue = 0.02, Bonferroni correction) (Keseler et al., 2013). Visual analysis of the genes whose expression was changed by inactivation of the *rsmF* gene revealed the possible involvement of Crp and σS, although the statistical support is insufficient (Keseler et al., 2013), and further partial verification with the help of RT qPCR (**Figure S5**) could not support the possible involvement of Crp in the response to the *rsmF* gene knockout. Notably, σS is among the proteins whose amount was decreased more than twofold upon RsmF inactivation. This may suggest that increased transcription of σ70-dependent reporter genes in this knockout strain might to some extent be explained by reduced competition over the RNA polymerase core enzyme. GO term analysis of the proteins whose abundance decreased upon RsmF rRNA methyltransferase inactivation revealed a TCA cycle (pvalue = 0.001). We found a

TABLE 2 | Significant differences in the proteome of *E. coli* caused by inactivation of *rlmC, rlmE*, and *rsmF* genes as revealed by panoramic proteome analysis.

Δ*rlmC*		Δ*rlmE*		Δ*rsmF*	
sra	0,28	gadB	0,21	astC	0,07
gadB	0,39	sra	0,23	acs	0,10
can	0,44	hyaB	0,30	modA	0,22
ytfE	2,05	cydA	0,31	pspE	0,22
		ybgS	0,40	ugpB	0,23
		appA	0,42	argT	0,25
		narG	0,44	sra	0,25
		mdtE	0,45	grcA	0,26
		glpQ	0,47	sdhA	0,27
		rpsU	0,48	msrB	0,28
		psiF	0,49	sdhB	0,30
		rpsL	0,50	aldA	0,31
		can	0,50	fdoH	0,32
		ansB	0,50	flu	0,33
		fadI	2,07	sthA	0,33
		nlpA	2,07	gadB	0,34
		fumA	2,14	gadC	0,34
		yfeX	2,14	sucB	0,35
		mdaB	2,21	yciF	0,36
		carA	2,21	psiF	0,36
		pck	2,26	putA	0,37
		carB	2,58	dadX	0,37
		metK	2,92	ynfK	0,37
		pyrI	3,22	osmY	0,39
		ydeN	3,32	ydfZ	0,39
		nanA	3,41	yhhA	0,41
		gdhA	3,52	yciE	0,41
		oppA	5,47	acnA	0,42
		ompF	6,86	rpoS	0,43
				yaiE	0,44
				frdB	0,44
				ybgS	0,44
				ynjE	0,45
				ompX	0,45
				tsx	0,46
				sucA	0,46
				hisJ	0,46
				iadA	0,47
				glcB	0,47
				mglB	0,48
				ftnA	0,48
				ybeL	0,49
				bfr	0,49
				sodC	0,49
				glnH	0,50
				nanA	2,03
				gatY	2,07
				miaB	2,09
				guaB	2,18
				iscS	2,19
				cysH	2,25
				metK	2,32
				cydA	2,35
				yidB	2,99
				ybeD	3,51
				mdaB	3,58
				oppA	3,61

Only proteins whose amount was changed more than twofold are listed. Complete analysis is available as ***Table S1****.*

certain degree of similarity between proteomic responses to *rlmC*, *rlmE*, and *rsmF* gene inactivation (**Table 2**). Interestingly, all three strains have a reduced concentration of Sra and GadB proteins.

Phylogenetic Distribution of Orthologs of *E. coli* rRNA Methyltransferases

Conservation of a protein may be a proxy of its functional importance. To compare the phylogenetic distribution of the complete set of *E. coli* rRNA methyltransferases, we performed blast searching (Altschul et al., 1990) for their orthologs. We aimed to determine a minimal phylogenetic group still containing orthologs of a particular *E. coli* rRNA methyltransferase. To this end, we verified whether any orthologs of a given *E. coli* rRNA methyltransferase might be identified outside Enterobacteriales, gamma Proteobacteria, Proteobacteria, and Bacteria, and whether any orthologs could be found in Eukarya and Archaea (**Figure 7**). We intentionally decided not to determine the entire phylogenetic trees of all rRNA methyltransferase families as this might be the matter of a separate, more focused, study.

The range of phylogenetic distributions in our set varied from universality for RsmA(KsgA) and RlmE, to extreme specificity to Enterobacteria for RsmJ and RlmG. Most frequently rRNA methyltransferase orthologs were found to be conserved among bacteria. For several *E. coli* rRNA methyltransferases, close paralogs were identified, such as RlmCD methyltransferases of gram positive bacteria for RlmC and RlmD (Desmolaize et al., 2011) or orthologs of RsmB in a BLAST search for RsmF homologues.

DISCUSSION

Several roles have been suggested for methylated rRNA nucleotides (Sergiev et al., 2011; Sergiev et al., 2018), among which ribosomal subunit assembly control appears amongst the most obvious ones. RsmA(KsgA) methyltransferase was shown to fulfill quality control over the last stages of small subunit assembly (Connolly et al., 2008), in line with a similar role of its eukaryotic homologue, Dim1, in ribosome biogenesis (Lafontaine et al., 1995). The mechanism of RsmA(KsgA)'s action suggested by Connolly et al. (2008) involved binding of the small subunit assembly intermediates by RsmA (KsgA), while its methyltransferase activity is delayed until the late stages of assembly. This type of activity might require more RsmA (KsgA) molecules relative to other rRNA MTs, which might cycle faster. Accordingly, we see the highest level of RsmA(KsgA) mRNA relative to mRNA coding for other rRNA MTs. We did not observe significant accumulation of assembly intermediates in the Δ*rsmA* (Δ*ksgA*) strain by sucrose gradient centrifugation, in line with earlier studies (Connolly et al., 2008). However, in consort with that work, we found an increase in the amount of the 17S rRNA precursor upon growth of the Δ*rsmA*(Δ*ksgA*) strain at a low temperature (**Figure 3**). Ubiquitous phylogenetic distribution of KsgA orthologs (**Figure 7A**) is explained by their involvement in the assembly of bacterial (Connolly et al., 2008) and eukaryal cytoplasmic (Lafontaine et al., 1995), as well as mitochondrial (Metodiev et al., 2009), small ribosomal subunits.

In addition, a significant amount of the 17S rRNA precursor accumulates in the Δ*rsmB* strain (**Figure 3**). RsmB is known to act on the early assembly intermediate of the 30S ribosome subunit, prior to the incorporation of the S19 protein (Weitzmann et al., 1991). It might be suggested that RsmB accelerates the assembly or prevents misassembly. Orthologs of RsmB are distributed among bacteria (**Figure 7A**), while in archaea and eukarya the same small subunit rRNA loop is modified to form acp^3U (Kowalak et al., 2000) or $m^1acp^3\psi$ (Wurm et al., 2010; Meyer et al., 2016) respectively. Enzymes Nep1 (Wurm et al., 2010) and Tsr3 (Meyer et al., 2016), responsible for the formation of the $m^1acp^3\psi$ nucleotide, are documented to serve a role in ribosome assembly and might be hypothesized to have a function reminiscent to that of RsmB.

Unexpectedly, inactivation of the 23S rRNA methyltransferase RlmC, responsible for the formation of m^5U747, also resulted in the accumulation of the 17S rRNA precursor (**Figure 3**). This phenomenon is not unique. Inactivation of the large subunit pseudouridine synthase RluD also resulted in the accumulation of the 17S rRNA precursor (Gutgsell et al., 2005), although later this phenotype was found to be dependent on the combination of the *rluD* knockout and *E. coli* K12-specific RF2 allele carrying threonine at the position 246 (O'Connor and Gregory, 2011). It seems likely that an influence of 23S rRNA methylation by RlmC on the processing of the small subunit rRNA is indirect, similar to that of RluD.

A severe assembly defect was known to be associated with inactivation of the RlmE methyltransferase (Bugl et al., 2000; Caldas et al., 2000; Hager et al., 2002; Arai et al., 2015). Here we corroborated and extended previous studies (**Figure 2**), demonstrating that while *rlmE* inactivation leads to a severe manifestation of ribosome misassembly, other rRNA MT knockouts have a marginal, if any, effect on the accumulation of ribosome assembly intermediates which might be distinguished by sucrose density centrifugation. In line with this, *rlmE* inactivation leads to the most significant growth retardation among the other rRNA MT gene knockouts. The severe phenotype of *rlmE* gene inactivation goes in parallel with the universal phylogenetic distribution of its orthologs (**Figure 7B**). A number of genes whose expression is dependent on σ^S and ppGpp, as well as ArcA and FNR, were found downregulated in the Δ*rlmE* strain (**Table 2**). This result may indicate lagging of the growth phase of the Δ*rlmE* strain relative to that of the wild type, which may also result in higher oxygenation of the bacterial culture sensed by ArcA and FNR. RlmC inactivation, which moderately affected the growth rate of bacteria, resulted in decreased expression of *sra* and *gadB* genes. Both these genes were similarly downregulated in the strains deficient in the *rlmE* and *rsmF* genes. This phenomenon might also reflect a lag in the growth phase of rRNA MT knockout cells, since both *sra* and *gadB* genes are normally expressed in the stationary phase (Castanie-Cornet and Foster, 2001; Izutsu et al., 2001).

Another common assumption is that rRNA modification might be functional in a particular set of conditions. In such a

FIGURE 7 | Phylogenetic distribution of the *E. coli* 16S rRNA **(A)** and 23S rRNA **(B)** methyltransferases orthologs. Shown are the simplified phylogenetic trees illustrating the occurrence of rRNA MT orthologs in the taxons beyond Enterobacteriales, gamma Proteobacteria, Proteobacteria, and Bacteria as well as in Eukarya and Archaea. Schematic trees are labelled by protein designations. Branches colored green correspond to those that contain orthologs of the methyltransferase, while those colored red do not.

case, it might be anticipated that the expression of the corresponding rRNA MT genes might be regulated in a condition-dependent manner. Analysis of rRNA MT gene expression throughout the bacterial growth cycle revealed that the predominant majority of rRNA MT genes is significantly upregulated in the exponential growth phase (**Figure 1**), along with the expression of rRNA operons. These data generally support a previous work (Sergiev et al., 2012) performed on the basis of GEO database analysis. Among the rRNA MT genes, *rlmE* and *rsmB* seem to be induced at the later growth phase, demonstrating a distinct way of transcriptional control. *RlmE* is known to be under heat shock control (Bugl et al., 2000), while

the mechanism of *rsmB* gene regulation remains unknown. Expression of the predominant majority of rRNA MT in the exponential growth phase does not support a function of methylated rRNA nucleotides within specific conditions.

Protein biosynthesis is one of the major resource consuming processes. Suboptimal protein biosynthesis efficiency caused by a lack of rRNA methylation may become a significant problem if resources are additionally spent on the synthesis of a protein coded in an artificially expressed gene. In line with this assumption, the growth rates of the majority of rRNA MT knockout strains are almost indistinguishable from that of the wild-type strain, except for the *rlmE* knockout, leading to the accumulation of ribosome assembly intermediates. However, induction of an exogenous gene, exemplified in this study by *rfp* and *cer*, decreases the growth rates of the rRNA MT knockout strains more significantly than that of the parental strain containing a complete inventory of rRNA MT coding genes. Almost all rRNA MT knockout strains were able to support only ½ to ¼ of the wild-type expression level of the constitutively expressed *rfp* gene. Additionally, induction of *cer* gene expression in the majority of the rRNA MT knockout strains not only failed to reach the wild-type expression levels, but also led to a more significant decrease in the expression of the *rfp* gene. This tendency is well reproduced by the evaluation of protein biosynthesis efficiency at a single cell level by flow cytometry. The likely interpretation of this fact is a mildly reduced protein synthesis capacity of the strains devoid of rRNA MT genes, which becomes a significant problem if a cell is under the burden of the synthesis of an exogenous protein. Among the rRNA MT deletion strains, the most compromised in the ability to synthesize exogenous proteins are the Δ*rlmA*, Δ*rlmG*, Δ*rlmH*, Δ*rlmI*, Δ*rlmJ*, Δ*rlmM*, Δ*rlmN*, Δ*rlmE,* Δ*rsmA* (Δ*ksgA*), Δ*rsmH,* and Δ*rsmJ* strains.

Apart from the general tendency, *rsmF* gene knockout resulted in a moderate increase of exogenous gene expression compared to the wild type. RsmF methyltransferase is responsible for the modification of m^5C1407 of the 16S rRNA (Andersen and Douthwaite, 2006), which increases the aminoglycoside sensitivity of *E. coli* (Gutierrez et al., 2012). The molecular mechanism of a small increase in exogenous gene expression in the Δ*rsmF* strain is unknown; a likely explanation is the suboptimal functioning of some regulatory mechanism. The transcriptional upregulation of reporter gene expression might be involved, at least partially. According to the proteome analysis, *rsmF* knockout leads to an alteration in the abundance of a large number of proteins, including decreased amounts of σ^S and downregulation of a number of genes transcribed with the help of this sigma factor. Perhaps, the increased propensity of the *rsmF* knockout strain to synthesize an exogenous protein might reflect a secondary effect.

This work might be of particular interest, as it is the first comprehensive study of all rRNA methyltransferase knockouts. Although only a few rRNA MT knockouts were found to have an influence on bacterial growth, ribosome assembly, or proteome, the majority of the knockout strains demonstrated a suboptimal capability to synthesize exogenous proteins. This may indicate a necessity of rRNA methylation at increased loads on the protein biosynthesis apparatus. More methods are needed to assess the subtle advantages of rRNA methylation in bacteria.

MATERIALS AND METHODS

Homology search was done with BLAST (Altschul et al., 1990), using the model organisms database. The resulting phylogenetic distribution of the BLAST hits was used to deduce whether any orthologous proteins could be found beyond Enterobacteriales, gamma Proteobacteria, Proteobacteria, and Bacteria, and whether any orthologs could be found in Eukarya and Archaea. Visual inspection of the search results was used to filter out paralogous, rather than orthologous proteins. In doubtful cases, a reciprocal BLAST search, with a putative orthologous protein identified in the original search, was performed to check whether the protein used as bait for the initial search would be found as its closest homologue among the proteins of *E. coli*.

In all experiments, the strains *ΔrsmG (JW3718), ΔrsmD (JW3430), ΔrsmB (JW3250), ΔrsmC (JW4333), ΔrsmH (JW0080), ΔrsmF (JW5301), ΔrsmE (JW2913), ΔrsmJ (JW5672), ΔrsmA(ksgA) (JW0050), ΔrlmA (JW1811), ΔrlmC (JW2756), ΔrlmF (JW5107), ΔrlmG (JW5513), ΔrlmH (JW0631), ΔrlmD (JW0843), ΔrlmI (JW5898), ΔrlmJ (JW3466), ΔrlmK/L (JW0931), ΔrlmB (JW4138), ΔrlmM (JW2777), ΔrlmN (JW2501), and ΔrlmE (JW3146)* from the Keio collection (Baba et al., 2006) were used and compared with the parental wild-type strain BW25113 (Datsenko and Wanner, 2000).

For the rRNA MT expression analysis, an overnight culture of wild-type E. coli was diluted in triplicate in fresh LB media to A260 0.01 and grown at 37°C in a shaker. Aliquots of cells were removed at 1, 2, 3, 4, 5, 6, 7, and 56 h and used for total RNA purification with Trizol reagent (Invitrogen), followed by cDNA synthesis with either a Maxima First Strand cDNA Synthesis Kit for RT-qPCR (Thermo) with a random hexamer primer or a Superscript reverse transcriptase (Invitrogen). Quantitative PCR was performed by a Maxima Hot Start DNA polymerase (Thermo) in the presence of SYBR green. The following primers were used for amplification of indicated mRNAs: rsmA(ksgA) (CCCTTTTGCGGGTTAATGGC and ACGCTT CGGGCAAAACTTTC), rsmB (CTATGCCACCTGTTCG GTGT and CTGTTTCGCAAAGTTCGGCA), rsmC (GCGC ATAATCTGCCAGCATCand AAGAACAAACCGGA AGCCCA), *rsmD* (CGCAAAAAGGTACACCGCAT and CAG CCAGCCGTTATCTTCCA), *rsmE* (TGAGCAGTGTGGTC GTAACCand ACCGGTAACGGCAACGTATT), *rsmF*(CCG ATTTTCTCGGTTGGGGA and ATACCACTCCTCCG CTTCCT), *rsmG* (GGACGCACGATAGAGAGTGGand TTC GGTCCGCGATCCTAATG), *rsmH* (CTCACGTCTGATCCTC TCGC and CAATAGTCTTCGCAACGGCG), *rsmJ* (AAT TCCAGATGTTCCGGCGT and TGCCTTATCTGTTCTG GCGG), *rlmA* (AAATCAGCCCCTTCAGCTCC and CCGATA CCAGTATGGACGCC), *rlmB* (GCCAGGACGTCA GTATCAGG and AGGATCAGCAGGAACGGTTG), *rlmC*

(GGGCTTTGGTTTACACTGCG and AAACTGAG TGGAGTCCAGCG), *rlmD* (TCGACAATGTCACTGGAGCC and GATGTTCCCTGGGGCTATCG), *rlmE* (AGGTCG ACAACCGTCATTCC and AAAGGGGTTACGTTCCCGTG), *rlmF* (CATCACAGCCGCTACGATCTand AAGTCTAC GCTTTGCTCCCC), *rlmG* (AATGCCAGTGTTTTCGGCAC and ACGGGATTGATGAGTCGAGC), *rlmH* (TGCT CACCCTCTTTGTCGAG and TTTACCGAGTACCTGCG TCG), *rlmI* (ATCGCGATAAGTACGCAGCA and GAAG CGCTGGATATTGCACG), *rlmJ* (CAGTTAGGCAGCGAAC ATGC and CTGACCGCTACGGTTGAAGT), *rlmKL* (TTTG AAACGTCTGCTGCGTG and CAGGCCGTCGAGATCC ATAC), rlmM (CTTCAACACGCAGTTCACCG and CATTTGCCGCCAGAAGATCG), *rlmN* (ATGTCG ATGGCTTCACCCTG and TATCGATGCTGCCTGTGGTC), *oppA* (AATCGTTCTTGAACGCAGC and GATCAA CGTGAACTTCGTCC), *metK* (AGGCTGAAGTGCGTAA AAAC and GGGCAGAATTGGCTTGATG), *nanA* (GTGG TGTACAACATTCCAGC and CGAAGATTTCGTCGTAA CCG), *gatY* (TTTGCCATCGCTTTGATGTC and TGG CATACATCCCATGAGC), *guaB* (AAGACTTCCAGAAAGC GGAA and TTCACGGATACGTTGCAGTA), *mdaB* (CAGC GACTACGATGTCAAAG and TTTTCGACGGATCTTT GCG), *ybeD* (CAGGCGTTACCTGAGCTG and TTTGC CCAGTTCTTCATACAGT), *astC* (ATTGACGACTCTACCTG TGC and TCACGCCGTAGTGCATATAG), *acs* (GCAG TATTCCGCTGAAGAAA and GATCTTCGGCGTTCATC TCT), *modA* (GCCTGCGGATCTGTTTATTT and TTCAG CAGTGAAGTCCAGTT), *psp* (ATCGATGTTCGTGTTCC AGA and CCCATCTCGCTAAGGATCTC), *ugpB* (GCTGG ATCCAACTGGAAAAC and TTCATCCTTACGACCGACG), *argT* (TACCGATAAACGTCAGCAGG and CCTT TACTACGCCAGGTCTC), *sra* (AATCGAACCGTCAGGCAC and TTTTCAGCGGGGCGTTT), *cer* (TGAGCAAGGGCGAG GAGC and TGGTGCAGATGAACTTCAGG), *rfp* (GCTGATCAAGGAGAACATGC and AGGATGTCG AAGGCGAAGG). Quantification of expression was done by ΔΔCt method using 16S rRNA as a reference (gAgAATgTgCCT TCgggAAC and CCgCTggCAACAAAggATAA for MT gene expression analysis or CATTGACGTTACCCGCAGAAGAAG and CTACGAGACTCAAGCTTGCCAGTA for other gene expression analyses). To estimate the proportion of 17S rRNA processing, an intermediate RT qPCR approach was used. The following primer sequences were used for the 16S rRNA (GAAGAGTTTGATCATGGCTCAG and CCACTCGTCAGCAAAGAAG) and for the 17S rRNA processing intermediate (TCATTACGAAGTTTAATTC TTTGAGCG and GAAGAGTTTGATCATGGCTCAG). The proportion 17S/(16S + 17S) was calculated by normalization of the levels of the 5′-end-extended 17S transcript to the total amount of 16S and 17S transcripts.

To assess the accumulation of assembly intermediates, cells of rRNA MT knockout strains and a BW25113 strain (WT) and were grown in 500 ml of an LB medium at 37°C or 20°C to A_{600} 0.6, slowly cooled on ice, and harvested by centrifugation. Cells pellets were resuspended in a lysis buffer (20 mM HEPES-KOH pH 7.5, 4.5 mM $Mg(OAc)_2$, 150 mM NH_4Cl, 4 mM β-mercaptoethanol, 0.05 mM spermine, 2 mM spermidine buffer) and lysed by ultrasonication. After removal of cell debris, lysates containing approximately 1,200 pmol of ribosomes were applied to either a 10% to 30% sucrose gradient in a buffer 20 mM HEPES-KOH pH 7.5, 1 mM Mg $(OAc)_2$, 200 mM NH_4Cl, 4 mM β-mercaptoethanol, or a 10% to 40% sucrose gradient in a buffer 20 mM HEPES-KOH pH 7.5, 10 mM $Mg(OAc)_2$, 200 mM NH_4Cl, 4 mM β-mercaptoethanol. Ultracentrifugation was performed by an SW41Ti rotor at 19,000 rpm for 19 h followed by optical density monitoring at 260 nm.

To create a pRFPCERtet construct, a pRFPCER plasmid (Osterman et al., 2013) was digested with HindIII and SacII and ligated with pair of pre-annealed complementary oligonucleotides (TetR F 5′ AGCTTGGGAAATCATAA AAAATTATTTGCTTACTCTATCATTGATAGAGT TATAATAGCCGC-3′ and TetR R 5′-GGCTATTATAA CTCTATCAATGATAGAGTAAGCAAATAATTTTTT ATGATTTCCCA-3′), containing a T5 promoter with the TetR binding site. The obtained plasmid was digested with SacII and NdeI restriction enzymes and ligated with pair of pre-annealed complementary oligonucleotides (5′-CACACAACAAAGG AGGTAC and 5′-TAGTACCTCCTTTGTTGTGTGGC), containing a highly efficient ribosomal binding site. The resulted plasmid was used for further study as pRFPCERtet.

To monitor growth rates upon exogenous gene overexpression, and to evaluate protein synthesis efficiency, cells of rRNA MT knockout strains and the BW25113 strain (WT) were transformed with the plasmid pRFPCERtet. Overnight cultures of the transformants, in triplicate for each strain, were diluted by LB with or without anhydrotetracycline 0.2 ug/ml to A_{600} 0.01 in a 96 well plate. Cells were cultivated with continuous shaking at 37°C with automatic A_{600} monitoring every 30 min by a Janus workstation (Perkin Elmer). Growth rates of rRNA MT knockout strains, and the wild-type strain not transformed by any plasmid, were measured likewise.

For evaluation of CER and RFP protein synthesis efficiency, cells transformed by the plasmid pRFPCERtet were grown in triplicates for 18 h in 200 ul LB media at 37°C with or without anhydrotetracycline 0.2 ug/ml in a 96 deep well plate with continuous shaking. After incubation, the cells were centrifuged in a 96 well plate and washed twice with 0.9% NaCl. The fluorescence of the cells was measured by a Victor X5 plate reader (Perkin Elmer) at 430/486 nm for CER and 531/595 nm for RFP.

To determine *in vivo* protein synthesis efficiency, the cells of the wild-type and rRNA MT knockout strains were transformed by a plasmid encoding the FastFT protein under a control of an araBAD promoter (Subach et al., 2009). Cells grown in LB media with 10 mM arabinose at 37°C after 48 h were diluted 1:100 by a fresh LB media with 10 mM arabinose. An aliquot was taken at various time points; cells were isolated by centrifugation, washed

two times by sterile PBS, and analyzed by a fluorescently activated cell sorter BD FACSAria III at the wavelengths 405/460 nm and 555/610 nm.

Comparative proteome analysis using 2D PAGE was performed as described (Hoch et al., 2015). Not less than three independently grown cultures were used for each knockout strain.

Shotgun comparative proteome analysis was performed as described (Toprak et al., 2014; Osterman et al., 2015). Briefly, cells resuspended in 0,75% w/w RapiGest SF (Waters) were lysed by sonication. After debris removal, protein cysteine bonds were reduced with 10 mM dithiothreitol and alkylated with 30 mM iodoacetamide. Trypsin was added in a 1/50 w/w ratio trypsin/protein and incubated at 37°C overnight. To stop trypsinolysis, trifluoroacetic (TFA) acid was added to the final concentration of 0,5% v/v. Peptides were desalted and resuspended in 3% acetonitrile (ACN), 0.1% TFA, to a final concentration of 2 μg/μl. Mass spectrometry analysis was performed on a TripleTOF 5600+ mass-spectrometer with a NanoSpray III ion source (ABSciex, Canada) coupled to a NanoLC Ultra 2D+ nano-HPLC system (Eksigent). For protein identification,.wiff data files were analyzed with ProteinPilot 4.5 revision 1656 (ABSciex) using the Paragon 4.5.0.0 revision 1654 (ABSciex) search algorithm and a standard set of identification settings to search against SwissProt database, species *Escherichia coli*.

DATA AVAILABILITY STATEMENT

The proteomic data in this article is available at ProteomeXchange, accession: PXD017171.

AUTHOR CONTRIBUTIONS

SE contributed to the design of experiments, 2D protein gel electrophoresis, exogenous gene expression analysis with reporter construct, data analysis. VT, AP, and PP contributed to 2D protein gel electrophoresis and data analysis. PP, EG, AZ, MD, and AG contributed to sucrose gradient centrifugation, exogenous gene expression analysis with reporter construct and RT qPCR expression analysis. PP, EG, and AZ performed 17S precursor quantification. MR performed the flow cytometry experiments. IO contributed to exogenous gene expression analysis with reporter construct. MS, OP, and VG contributed to proteome analysis. AB, PS, and OD contributed to the design of the study. All authors contributed to the manuscript writing.

FUNDING

This work was funded by the Russian Science Foundation grant 19-14-00043 in a part related to growth rate measurement, proteome analysis and RT qPCR, Russian Foundation for Basic Research grant 17-00-00366 in a part related to reporter constructs. Scholarships to the authors and equipment were provided in part by Lomonosov Moscow State University's government funding and development program.

ACKNOWLEDGMENTS

We thank Dr. H. Mori, NIG, Japan for providing us with the knockout strains from the Keio collection. Authors are thankful to Alex Lebedeff for improvement of the English of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00097/full#supplementary-material

REFERENCES

Agarwalla, S., Kealey, J. T., Santi, D. V., and Stroud, R. M. (2002). Characterization of the 23S ribosomal RNA m5U1939 methyltransferase from Escherichia coli. *J. Biol. Chem.* 277, 8835–8840. doi: 10.1074/jbc.M111825200

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Andersen, N., and Douthwaite, S. (2006). YebU is a m5C methyltransferase specific for 16S rRNA nucleotide 1407. *J. Mol. Biol.* 359, 777–786. doi: 10.1016/j.jmb.2006.04.007

Arai, T., Ishiguro, K., Kimura, S., Sakaguchi, Y., Suzuki, T., and Suzuki, T. (2015). Single methylation of 23S rRNA triggers late steps of 50S ribosomal subunit assembly. *Proc. Natl. Acad. Sci. U. S. A.* 112, E4707–E4716. doi: 10.1073/pnas.1506749112

Bugl, H., Fauman, E., Staker, B., Zheng, F., Kushner, S., Saper, M., et al. (2000). RNA methylation under heat shock control. *Mol. Cell* 6, 349–360. doi: 10.1016/S1097-2765(00)00035-6

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., et al. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* 2, 2006.0008. doi: 10.1038/msb4100050

Basturea, G. N., Rudd, K. E., and Deutscher, M. P. (2006). Identification and characterization of RsmE, the founding member of a new RNA base methyltransferase family. *RNA* 12, 426–434. doi: 10.1261/rna.2283106

Basturea, G. N., Dague, D. R., Deutscher, M. P., and Rudd, K. E. (2012). YhiQ is RsmJ, the methyltransferase responsible for methylation of G1516 in 16S rRNA of *E. coli*. *J. Mol. Biol.* 415, 16–21. doi: 10.1016/j.jmb.2011.10.044

Burakovsky, D., Prokhorova, I., Sergiev, P., Milón, P., Sergeeva, O., Bogdanov, A., et al. (2012). Impact of methylations of m2G966/m5C967 in 16S rRNA on bacterial fitness and translation initiation. *Nucleic Acids Res.* 40, 7885–7895. doi: 10.1093/nar/gks508

Caldas, T., Binet, E., Bouloc, P., and Richarme, G. (2000). Translational defects of Escherichia coli mutants deficient in the Um(2552) 23S ribosomal RNA methyltransferase RrmJ/FTSJ. *Biochem. Biophys. Res. Commun.* 271, 714–718. doi: 10.1006/bbrc.2000.2702

Castanie-Cornet, M. P., and Foster, J. W. (2001). Escherichia coli acid resistance: cAMP receptor protein and a 20 bp cis-acting sequence control pH and stationary phase expression of the gadA and gadBC glutamate decarboxylase genes. *Microbiology* 147, 709–715. doi: 10.1099/00221287-147-3-709

Connolly, K., Rife, J. P., and Culver, G. (2008). Mechanistic insight into the ribosome biogenesis functions of the ancient protein KsgA. *Mol. Microbiol.* 70, 1062–1075. doi: 10.1111/j.1365-2958.2008.06485.x

Datsenko, K. A., and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* 97, 6640–6645. doi: 10.1073/pnas.120163297

Desmolaize, B., Fabret, C., Brégeon, D., Rose, S., Grosjean, H., and Douthwaite, S. (2011). A single methyltransferase YefA (RlmCD) catalyses both m5U747 and m5U1939 modifications in *Bacillus subtilis* 23S rRNA. *Nucleic Acids Res.* 39, 9368–9375. doi: 10.1093/nar/gkr626

Ero, R., Peil, L., Liiv, A., and Remme, J. (2008). Identification of pseudouridine methyltransferase in *Escherichia coli*. *RNA* 14, 2223–2233. doi: 10.1261/rna.1186608

Golovina, A., Dzama, M., Osterman, I., Sergiev, P., Serebryakova, M., Bogdanov, A., et al. (2012). The last rRNA methyltransferase of E. coli revealed: the yhiR gene encodes adenine-N6 methyltransferase specific for modification of A2030 of 23S ribosomal RNA. *RNA* 18, 1725–1734. doi: 10.1261/rna.034207.112

Green, R., and Noller, H. F. (1996). In vitro complementation analysis localizes 23S rRNA posttranscriptional modifications that are required for Escherichia coli 50S ribosomal subunit assembly and function. *RNA* 2, 1011–1021.

Gu, X. R., Gustafsson, C., Ku, J., Yu, M., and Santi, D. V. (1999). Identification of the 16S rRNA m^5C967 methyltransferase from *Escherichia coli*. *Biochemistry* 38, 4053–4057. doi: 10.1021/bi982364y

Gustafsson, C., and Persson, B. C. (1998). Identification of the *rrmA* gene encoding the 23S rRNA m1G745 methyltransferase in *Escherichia coli* and characterization of an m^1G745-deficient mutant. *J. Bacteriol.* 180, 359–365. doi: 10.1128/JB.180.2.359-365.1998

Gutgsell, N. S., Deutscher, M. P., and Ofengand, J. (2005). The pseudouridine synthase RluD is required for normal ribosome assembly and function in *Escherichia coli*. *RNA*. 11, 1141–1152. doi: 10.1261/rna.2550105

Gutierrez, B., Escudero, J., San Millan, A., Hidalgo, L., Carrilero, L., Ovejero, C., et al. (2012). Fitness cost and interference of Arm/Rmt aminoglycoside resistance with the RsmF housekeeping methyltransferases. *Antimicrob. Agents Chemother.* 56, 2335–2341. doi: 10.1128/AAC.06066-11

Hager, J., Staker, B., Bugl, H., and Jakob, U. (2002). Active site in RrmJ, a heat shock-induced methyltransferase. *J. Biol. Chem.* 277, 41978–41986. doi: 10.1074/jbc.M205423200

Helser, T. L., Davies, J. E., and Dahlberg, J. E. (1972). Mechanism of kasugamycin resistance in *Escherichia coli*. *Nat. New Biol.* 235, 6–9. doi: 10.1038/newbio235006a0

Hoch, P. G., Burenina, O. Y., Weber, M. H., Elkina, D. A., Nesterchuk, M. V., Sergiev, P. V., et al. (2015). Phenotypic characterization and complementation analysis of Bacillus subtilis 6S RNA single and double deletion mutants. *Biochimie* 117, 87–99. doi: 10.1016/j.biochi.2014.12.019

Izutsu, K., Wada, C., Komine, Y., Sako, T., Ueguchi, C., Nakura, S., et al. (2001). Escherichia coli ribosome-associated protein SRA, whose copy number increases during stationary phase. *J. Bacteriol.* 183, 2765–2773. doi: 10.1128/JB.183.9.2765-2773.2001

Keseler, I., Mackie, A., Peralta-Gil, M., Santos-Zavaleta, A., Gama-Castro, S., Bonavides-Martínez, C., et al. (2013). EcoCyc: fusing model organism databases with systems biology. *Nucleic Acids Res.* 41, D605–D612. doi: 10.1093/nar/gks1027

Kimura, S., and Suzuki, T. (2010). Fine-tuning of the ribosomal decoding center by conserved methyl- modifications in the Escherichia coli 16S rRNA. *Nucleic Acids Res.* 38, 1341–1352. doi: 10.1093/nar/gkp1073

Kimura, S., Ikeuchi, Y., Kitahara, K., Sakaguchi, Y., Suzuki, T., and Suzuki, T. (2012). Basemethylations in the double-stranded RNA by a fused methyltransferase bearing unwinding activity. *Nucleic Acids Res.* 40, 4071–4085. doi: 10.1093/nar/gkr1287

Kowalak, J. A., Bruenger, E., Crain, P. F., and McCloskey, J. A. (2000). Identities and phylogenetic comparisons of posttranscriptional modifications in 16S ribosomal RNA from *Haloferax volcanii*. *J. Biol. Chem.* 275, 24484–24489. doi: 10.1074/jbc.M002153200

Lafontaine, D., Vandenhaute, J., and Tollervey, D. (1995). The 18S rRNA dimethylase Dim1p is required for pre-ribosomal RNA processing in yeast. *Genes Dev.* 9, 2470–2481. doi: 10.1101/gad.9.20.2470

Lesnyak, D. V., Osipiuk, J., Skarina, T., Sergiev, P. V., Bogdanov, A. A., Edwards, A., et al. (2007). Methyltransferase that modifies guanine 966 of the 16S rRNA: functional identification and tertiary structure. *J. Biol. Chem.* 282, 5880–5887. doi: 10.1074/jbc.M608214200

Li, G., Burkhardt, D., Gross, C., and Weissman, J. (2014). Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell* 157, 624–635. doi: 10.1016/j.cell.2014.02.033

Lovgren, J. M., and Wikstrom, P. M. (2001). The *rlmB* gene is essential for formation of Gm2251 in 23S rRNA but not for ribosome maturation in *Escherichia coli*. *J. Bacteriol.* 183, 6957–6960. doi: 10.1128/JB.183.23.6957-6960.2001

Madsen, C. T., Mengel-Jorgensen, J., Kirpekar, F., and Douthwaite, S. (2003). Identifying the methyltransferases for m5U747 and m5U1939 in 23S rRNA using MALDI mass spectrometry. *Nucleic Acids Res.* 31, 4738–4746. doi: 10.1093/nar/gkg657

Metodiev, M. D., Lesko, N., Park, C. B., Cámara, Y., Shi, Y., Wibom, R., et al. (2009). Methylation of 12S rRNA is necessary for *in vivo* stability of the small subunit of the mammalian mitochondrial ribosome. *Cell Metab.* 9, 386–397. doi: 10.1016/j.cmet.2009.03.001

Meyer, B., Wurm, J. P., Sharma, S., Immer, C., Pogoryelov, D., Kötter, P., et al. (2016). Ribosome biogenesis factor Tsr3 is the aminocarboxypropyl transferase responsible for 18S rRNA hypermodification in yeast and humans. *Nucleic Acids Res.* 44, 4304–4316. doi: 10.1093/nar/gkw244

O'Connor, M., and Gregory, S. T. (2011). Inactivation of the RluD pseudouridine synthase has minimal effects on growth and ribosome function in wild-type *Escherichia coli* and *Salmonella enterica*. *J. Bacteriol.* 193, 154–162. doi: 10.1128/JB.00970-10

Okamoto, S., Tamaru, A., Nakajima, C., Nishimura, K., Tanaka, Y., Tokuyama, S., et al. (2007). Loss of a conserved 7-methylguanosine modification in 16S rRNA confers low-level streptomycin resistance in bacteria. *Mol. Microbiol.* 63, 1096–1106. doi: 10.1111/j.1365-2958.2006.05585.x

Osterman, I. A., Evfratov, S. A., Sergiev, P. V., and Dontsova, O. A. (2013). Comparison of mRNA features affecting translation initiation and reinitiation. *Nucleic Acids Res.* 41, 474–486. doi: 10.1093/nar/gks989

Osterman, I. A., Evfratov, S. A., Dzama, M. M., Pletnev, P. I., Kovalchuk, S. I., Butenko, I. O., et al. (2015). A bacterial homolog YciH of eukaryotic translation initiation factor eIF1 regulates stress-related gene expression and is unlikely to be involved in translation initiation fidelity. *RNA Biol.* 12, 966–971. doi: 10.1080/15476286.2015.1069464

Poldermans, B., Roza, L., and Van Knippenberg, P. H. (1979). Studies on the function of two adjacent N6,N6-dimethyladenosines near the 3' end of 16S ribosomal RNA of *Escherichia coli*. III. Purification and properties of the methylating enzyme and methylase-30S interactions. *J. Biol. Chem.* 254, 9094–9100.

Prokhorova, I., Osterman, I., Burakovsky, D., Serebryakova, M., Galyamina, M., Pobeguts, O., et al. (2013). Modified nucleotides m(2)G966/m(5)C967 of *Escherichia coli* 16S rRNA are required for attenuation of tryptophan operon. *Sci. Rep.* 3, 3236. doi: 10.1038/srep03236

Purta, E., Kaminska, K. H., Kasprzak, J. M., Bujnicki, J. M., and Douthwaite, S. (2008a). YbeA is the $m^3\Psi$ methyltransferase RlmH that targets nucleotide 1915 in 23S rRNA. *RNA* 14, 2234–2244. doi: 10.1261/rna.1198108

Purta, E., O'Connor, M., Bujnicki, J., and Douthwaite, S. (2008b). YccW is the m^5C methyltransferase specific for 23S rRNA nucleotide 1962. *J. Mol. Biol.* 383, 641–651. doi: 10.1016/j.jmb.2008.08.061

Purta, E., O'Connor, M., Bujnicki, J. M., and Douthwaite, S. (2009). YgdE is the 2'-O-ribose methyltransferase RlmM specific for nucleotide C2498 in bacterial 23S rRNA. *Mol. Microbiol.* 72, 1147–1158. doi: 10.1111/j.1365-2958.2009.06709.x

Sergiev, P. V., Lesnyak, D. V., Bogdanov, A. A., and Dontsova, O. A. (2006). Identification of Escherichia coli m(2)G methyltransferases: II. The ygjO gene encodes a methyltransferase specific for G1835 of the 23 S rRNA. *J. Mol. Biol.* 364, 26–31. doi: 10.1016/j.jmb.2006.09.008

Sergiev, P. V., Serebryakova, M. V., Bogdanov, A. A., and Dontsova, O. A. (2008). The *ybiN* gene of *Escherichia coli* encodes adenine-N6 methyltransferase specific for modification of A1618 of 23S ribosomal RNA, a methylated residue located close to the ribosomal exit tunnel. *J. Mol. Biol.* 375, 291–300. doi: 10.1016/j.jmb.2007.10.051

Sergiev, P., Golovina, A., Prokhorova, I., Sergeeva, O., Osterman, I., Nesterchuk, M., et al. (2011). "Modifications of ribosomal RNA: from enzymes to function," in *Ribosomes: Structure, Function, and Dynamics*. Eds. M. V. Rodnina, W. Wintermeyer and R. Green (Wien, New York: Springer), 97–110.

Sergiev, P., Golovina, A., Sergeeva, O., Osterman, I., Nesterchuk, M., Bogdanov, A., et al. (2012). How much can we learn about the function of bacterial rRNA

modification by mining large-scale experimental datasets? *Nucleic Acids Res.* 40, 5694–5705. doi: 10.1093/nar/gks219

Sergiev, P., Aleksashin, N., Chugunova, A., Polikanov, Y., and Dontsova, O. (2018). Structural and evolutionary insights into ribosomal RNA methylation. *Nat. Chem. Biol.* 14, 226–235. doi: 10.1038/nchembio.2569

Shajani, Z., Sykes, M., and Williamson, J. (2011). Assembly of bacterial ribosomes. *Annual. Rev. Biochem.* 80, 501–526. doi: 10.1146/annurev-biochem-062608-160432

Smith, B. A., Gupta, N., Denny, K., and Culver, G. M. (2018). Characterization of 16S rRNA processing with pre-30S subunit assembly intermediates from. *E. Coli. J. Mol. Biol.* 430, 1745–1759. doi: 10.1016/j.jmb.2018.04.009

Subach, F. V., Subach, O. M., Gundorov, I. S., Morozova, K. S., Piatkevich, K. D., Cuervo, A. M., et al. (2009). Monomeric fluorescent timers that change color from blue to red report on cellular trafficking. *Nat. Chem. Biol.* 5, 118–126. doi: 10.1038/nchembio.138

Tan, J., Jakob, U. 0., and Bardwell, J. (2002). Overexpression of two different GTPases rescues a null mutation in a heat-induced rRNA methyltransferase. *J. Bact.* 184, 2692–2698. doi: 10.1128/JB.184.10.2692-2698.2002

Toh, S. M., and Mankin, A. S. (2008). An indigenous posttranscriptional modification in the ribosomal peptidyl transferase center confers resistance to an array of protein synthesis inhibitors. *J. Mol. Biol.* 380, 593–597. doi: 10.1016/j.jmb.2008.05.027

Toprak, U. H., Gillet, L. C., Maiolica, A., Navarro, P., Leitner, A., and Aebersold, R. (2014). Conserved peptide fragmentation as a benchmarking tool for mass spectrometers and a discriminating feature for targeted proteomics. *Mol. Cell. Proteomics* 13, 2056–2071. doi: 10.1074/mcp.O113.036475

Tscherne, J. S., Nurse, K., Popienick, P., Michel, H., Sochacki, M., and Ofengand, J. (1999a). Purification, cloning, and characterization of the 16S RNA m5C967 methyltransferase from *Escherichia coli*. *Biochemistry* 38, 1884–1892. doi: 10.1021/bi981880l

Tscherne, J. S., Nurse, K., Popienick, P., and Ofengand, J. (1999b). Purification, cloning, and characterization of the 16S RNA m2G1207 methyltransferase from Escherichia coli. *J. Biol. Chem.* 274, 924–929. doi: 10.1074/jbc.274.2.924

Vazquez-Laslop, N., Thum, C., and Mankin, A. S. (2008). Molecular mechanism of drug-dependent ribosome stalling. *Mol. Cell* 30, 190–202. doi: 10.1016/j.molcel.2008.02.026

Vazquez-Laslop, N., Ramu, H., Klepacki, D., Kannan, K., and Mankin, A. (2010). The key function of a conserved and modified rRNA residue in the ribosomal response to the nascent peptide. *EMBO J.* 29, 3108–3117. doi: 10.1038/emboj.2010.180

Weisblum, B. (1995). Erythromycin resistance by ribosome modification. *Antimicrob. Agents Chemother.* 39, 577–585. doi: 10.1128/AAC.39.3.577

Weitzmann, C., Tumminia, S. J., Boublik, M., and Ofengand, J. (1991). A paradigm for local conformational control of function in the ribosome: binding of ribosomal protein S19 to Escherichia coli 16S rRNA in the presence of S7 is required for methylation of m^2G966 and blocks methylation of m^5C967 by their respective methyltransferases. *Nucleic Acids Res.* 19, 7089–7095. doi: 10.1093/nar/19.25.7089

Wurm, J. P., Meyer, B., Bahr, U., Held, M., Frolow, O., Kötter, P., et al. (2010). The ribosome assembly factor Nep1 responsible for Bowen-Conradi syndrome is a pseudouridine-N1-specific methyltransferase. *Nucleic Acids Res.* 38, 2387–2398. doi: 10.1093/nar/gkp1189

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

ORIGINAL RESEARCH
published: 28 February 2020
doi: 10.3389/fgene.2020.00166

Upstream ORFs Influence Translation Efficiency in the Parasite *Trypanosoma cruzi*

*Santiago Radío[1,2], Beatriz Garat[2], José Sotelo-Silveira[1,3] and Pablo Smircich[1,2]**

[1] Department of Genomics, Instituto de Investigaciones Biológicas Clemente Estable, Ministerio de Educación y Cultura, Montevideo, Uruguay, [2] Laboratory of Molecular Interactions, Facultad de Ciencias, Universidad de la República, Montevideo, Uruguay, [3] Department of Cell and Molecular Biology, Facultad de Ciencias, Universidad de la República, Montevideo, Uruguay

It is generally accepted that the presence of ORFs in the 5′ untranslated region of eukaryotic transcripts modulates the production of proteins by controlling the translation initiation rate of the main CDS. In trypanosomatid parasites, which almost exclusively depend on post-transcriptional mechanisms to regulate gene expression, translation has been identified as a key step. However, the mechanisms of control of translation are not fully understood. In the present work, we have annotated the 5′UTRs of the *Trypanosoma cruzi* genome both in epimastigotes and metacyclic trypomastigotes and, using a stringent classification approach, we identified putative regulatory uORFs in about 9% of the analyzed 5′UTRs. The translation efficiency (TE) and translational levels of transcripts containing putative repressive uORFs were found to be significantly reduced. These findings are supported by the fact that proteomic methods only identify a low number of proteins coded by transcripts containing repressive uORF. We additionally show that AUG is the main translation initiator codon of repressive uORFs in *T. cruzi*. Interestingly, the decrease in TE is more pronounced when the uORFs overlaps the main CDS. In conclusion, we show that the presence of the uORF and features such as initiation codon and/or location of the uORFs may be acting to fine tune translation levels in these parasites.

Keywords: *Trypanosoma cruzi*, uORF, 5′UTR, translation efficiency, translation regulation

OPEN ACCESS

Edited by:
Assaf Katz,
University of Chile, Chile

Reviewed by:
Gary Loughran,
University College Cork, Ireland
Michael James Holmes,
Indiana University Bloomington, United States

***Correspondence:**
Pablo Smircich
psmircich@fcien.edu.uy

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *06 December 2019*
Accepted: *12 February 2020*
Published: *28 February 2020*

Citation:
Radío S, Garat B, Sotelo-Silveira J and Smircich P (2020) Upstream ORFs Influence Translation Efficiency in the Parasite Trypanosoma cruzi. Front. Genet. 11:166. doi: 10.3389/fgene.2020.00166

INTRODUCTION

Translation regulation depends on signals that are present mainly at the untranslated regions of the mRNAs (UTRs). Although many regulatory elements are present in the 3′UTR regions, 5′UTRs may also contain important *cis* acting regulators. In particular, the presence of small open reading frames (upstream ORFs, uORF) in this region has been described as a regulatory mechanism influencing the formation of the translation initiation complex at the initiator codon of the main CDS (McCarthy, 1998). Typically, the presence of uORFs decreases the efficiency of initiation at the main CDS, thus leading to downregulation of its translation rate (Griffin et al., 2001; Vattem and Wek, 2004; Chen et al., 2010). Even though there are cases where the presence of a uORF positively influences the translation efficiency (TE) of the main CDS, these are exceptions (Griffin et al., 2001; Vattem and Wek, 2004; Chen et al., 2010). uORFs may serve as elements that respond to altered environmental conditions, allowing the cells to rapidly adjust their protein production rates (Calvo et al., 2009; Lawless et al., 2009). Features of the uORFs that are efficiently translated

and negatively regulate the main CDS have been defined. Among them, the uORF position within the 5′UTR, its length, the sequence context of the initiation codon and the overlap with the main CDS have all been shown to affect their regulatory potential (Kozak, 1987, 2002; Rajkowitsch et al., 2004; Wethmar, 2014; Chew et al., 2016; Fervers et al., 2018).

Kinetoplastids are excellent models for the analysis of post-transcriptional mechanisms of gene expression regulation, given that transcription initiation is considered to be constitutive for most of the genome (Soldatos et al., 2010). This implies that the control of the substantial and rapid cell biology changes that take place throughout their life cycle, must rely mostly on post-transcriptional mechanisms (Soldatos et al., 2010). In addition to studies on regulation of mRNA steady state levels and protein degradation or modification, ribosome profiling approaches (Ingolia et al., 2011; Ivanov et al., 2011; Smircich et al., 2015) have confirmed that translation is a key step to regulate gene expression profiles in these organisms (Peabody, 1989; Vasquez et al., 2014; Smircich et al., 2015).

The role of uORFs in kinetoplastids has been studied using reporter genes and through the analysis of Ribo-Seq and proteomic data (Siegel et al., 2005; Jensen et al., 2014; Vasquez et al., 2014; Fervers et al., 2018). Probably due to the use of different definitions, the number of CDS associated to uORFs in the kinetoplastid genomes remains controversial, ranging from 11–22% in *T. brucei* (according to (Jensen et al., 2014) or Siegel et al., 2005, respectively) or 29% in the *T. congolense* (according to Fervers et al., 2018).

Except for one early report about uORFs composition in four specific genes (Jaeger and Brandao, 2011), a systematic analysis of uORFs in *Trypanosoma cruzi* is missing. This parasitic protozoan, is the causative agent of the neglected Chagas' disease a neglected infection affecting millions of people in Latin America (Lidani et al., 2019). It has a complex life cycle involving replicative and non-replicative developmental forms both in vertebrate (amastigote and bloodstream trypomastigotes, respectively) and invertebrate hosts (epimastigotes and metacyclic trypomastigotes, respectively) (Tyler and Engman, 2001).

In the present study, we performed a systematic search for uORFs in *T. cruzi* 5′UTRs focusing on those with repressive potential. The two life cycle forms of the parasite in the invertebrate host (epimastigotes and metacyclic trypomastigotes) were analyzed to detect putative stage-specific control. Using a stringent approach that considers factors that influence the uORF TE, we found that at least 5% of the 5′UTRs of the mRNAs in the *T. cruzi* genome contain repressive uORFs (approx. 9% of the analyzed 5′UTRs). Our results show that genes containing these uORFs have a low translational efficiency and low translation levels in both epimastigotes and metacyclic trypomastigotes insect stages. We also show that the AUG codon is mainly responsible for initiating translation of the uORFs that cause this effect. Additionally, we analyzed the uORFs that overlap the main CDS, a category not previously analyzed in *T. cruzi*. We found that these overlapping uORFs (uORFo) are associated to the most pronounced decrease of translation efficiency in this organism. In conclusion, we present a repertoire of genes in *T. cruzi* that exhibit putative repressive uORFs at the 5′UTR and define characteristics such as the identity of the initiation codon and/or location of the uORFs, that may be contributing to the regulation of gene expression by fine tuning the translation levels.

MATERIALS AND METHODS

Genomic Data

The genome sequence of *T. cruzi* strain CL Brener and its gene annotation were obtained from the TriTrypDB database (version 32).

UTRs Determination

The UTR sequences were determined for the epimastigote and metacyclic trypomastigote stages using the UTRme software (Radio et al., 2018), based on the transcriptomic data (accession numbers ranging from SRR1346053–SRR1346059) data generated in Li et al. (2016). The 5′UTRs with the highest scoring *trans*-spliced site and longer than 5 nucleotides were used. 8206 5′UTR regions were annotated for epimastigotes and 8217 for metacyclic trypomastigotes.

Multigene family members were not analyzed because, due to either the large content of repeated regions or to assembly problems, the 5′UTR ends are difficult to assess. Besides, their high UTR similarity and abundance would bias the results. Multigene families removed include large families of surface proteins such as MASPs, GP63, Mucins, and TcTS. Finally, we also decided to remove 5′UTR that contain fragments of *T. cruzi* protein coding regions, likely produced by assembly or annotation errors. For this purpose, the BLASTX tool (Altschul et al., 1990) was used, eliminating UTRs that returned a hit against the *T. cruzi* CDS with an *e*-value of less than 0.005.

uORF Annotation and Classification

Annotation of 5′UTR sequences was performed using the UTRme tool (Radio et al., 2018) using data obtained from Li et al. (2016). Open reading frames (ORF) greater than three codons were obtained using the getORF tool (Rice et al., 2000). For all ORFs present in the same strand as the CDS, AUGs in frame to the stop codon of the ORF (uAUG) were searched to define all putative coding sequences. uORFs in which the coding sequences start in the 5′UTR region and end within the main CDS (out of frame with respect to the main CDS AUG) were defined as overlapping uORFs (uORFo) while the ones contained entirely in the 5′UTR were defined as non-overlapping uORF (uORFno), in this case the frame with respect to the main CDS AUG was not considered.

Putative repressive uORFno elements were filtered considering the following characteristics: (1) the presence of an AUG start codon; (2) a minimum distance of 15 nucleotides from the uAUG to the 5′ end; (3) a maximum distance of 50 nucleotides between the stop codon of the uORF and the AUG of the main CDS; (4) minimum length of 5 amino acids. In the case of the uORFo the requirements were: (1) to start with AUG; (2) a minimum length of three amino acids before the AUG of the main CDS.

To identify and study uORFs with coding sequences initiating from near cognate codons (NCC), the annotation and classification strategy remained the same with the exception that the initiator codon was changed from AUG to the codon under study.

Translation Efficiency Determination

Ribosome Profiling data were obtained from the SRA: PRJNA260933 (Smircich et al., 2015) and correspond to the epimastigote and trypomastigote metacyclic stages of *T. cruzi*. The cutadapt program (Martin, 2011) was used to remove the adapters and filter by quality. The same parameters for the specification of the adapter (5′- CGCCTTGGCCGTACAGCAG - 3′), minimum quality allowed (13 phred score), maximum allowed error rate (0.1), and colorspace mode were used for both the RNA-Seq and Ribo-Seq data. As for the length limitation, a size larger than 18 base pairs (bp) and a range between 25 and 40 bp were defined for RNA-Seq and Ribo-Seq data, respectively.

The Bowtie aligner (version 1.2.2) (Langmead et al., 2009) was used to remove contamination produced by readings of ribosomal RNA origin. *T. cruzi* rRNAs were downloaded from the TriTrypDB database. ShortStack (version 3.6) (Johnson et al., 2016) was used to align the previously obtained reads. The ShortStack program was adapted to accept data from SOLiD technology, while the prediction of secondary structures and micro RNAs was disabled. The mapping mode chosen was the single weighting mode (U) where, only the frequencies of the uniquely aligned reads in the vicinity of the alignment in question are considered in the final weighting. Then, the FeatureCounts module of the SubRead package (v1.5.2) (Liao et al., 2013) was used to quantify the number of reads originated in each transcript or CDS (for RNA-seq or Ribo-seq data, respectively). Finally, the translation efficiency in each stage was obtained with the RiboDiff software (Zhong et al., 2017).

Generation of Sequence Logos

Sequence logos were generated with the WebLogo 3 online tool, using default parameters (Crooks et al., 2004).

Gene Ontology Enrichment Analysis

Gene ontology analysis were performed using the tool available for this purpose in the TriTrypDB database for this purpose. The visualization and reduction of the categories was carried out by REVIGO (Supek et al., 2011) in conjunction with the graphic environment of the R language (R Core Team, 2013).

RESULTS AND DISCUSSION

The *T. cruzi* Genome Contains Hundreds of uORFs With Repressive Potential

Precise definition of 5′UTR sequences was performed using the UTRme tool specifically developed to characterize untranslated regions of trypanosomatid genomes (Radio et al., 2018) using deep transcriptomic data obtained from Li et al. (2016). All open reading frames greater than three codons and present in the same strand as the CDS were then obtained. As described in the methods section, uORFs were classified as overlapping (uORFo) if the coding sequence starts in the 5′UTR region and ends within the main CDS (out of frame), or non-overlapping uORF (uORFno) which are contained entirely in the 5′UTR of the mRNA.

Since not all uORFs have coding potential nor do they have the same translation initiation efficiency, it is crucial to consider features that influence the repressive capacity of an uORF to define them. According to the literature, putative repressive uORFno elements are more likely to contain the following characteristics: (1) the presence of an AUG start codon (Clements et al., 1988); (2) a minimum distance of 15 nucleotides from the 5′ end, since shorter distances render difficult the assembly of the translation machinery at the uORF initiation codon (Vilela and McCarthy, 2003); (3) a maximum distance of 50 nucleotides between the stop codon of the uORF and the AUG of the main CDS, as shorter re-initiation times are associated to greater repressive potential (Chew et al., 2016) and (4) minimum length of 5 amino acids, as the longer the uORF length, the lower the probability of translation reinitiating at the main CDS (Kozak, 2001; Rajkowitsch et al., 2004). In the case of the uORFo, in which the coding sequences ends within the main CDS the requirements were: (1) to start with AUG; and (2) at least a minimum length of three amino acids before the AUG of the main CDS. Both types of uORFs passing the requirements were classified as repressive uORFs and 5′UTR regions that have one or more putative repressive elements were also classified as repressive. In turn, 5′UTRs that do not contain repressive elements, have a size greater than 50 nucleotides and do not contain AUGs were defined as non-repressive. Regions that do not fulfill any of the above categories were not assigned any classification.

After performing the above classification, the 5′UTR regions of *T. cruzi* mRNAs in both insect stages were studied. For the epimastigote stage, 6744 regions were analyzed, 568 of which (8.4%) classified as repressive. Among them, 111 regions contain only uORFo and 160 contain only uORFno, while the remaining regions (297) encompass elements of both categories. In addition, 3375 regions (50%) were classified as non-repressive. In the trypomastigote stage, 6750 mRNAs were analyzed, and a similar distribution was observed, 602 regions (8.9%) being classified as repressive. Among them, 231 carry both repressive types of uORFs, while 204 contain only uORFo and 167 only uORFno. Finally, 3333 regions (49%) were classified as non-repressive. These numbers are summarized in **Table 1** and the list of gene identifiers for the analyzed UTRs can be found in **Supplementary Table 1**. Interestingly, the percentage of genes that have putative repressive elements is similar to that detected in *T. brucei* (11%) by Jensen et al. (2014). The observed differences with other reports (Vasquez et al., 2014; Fervers et al., 2018) are most likely due to the different inclusion criteria applied in each work.

mRNAs Containing Putative Repressive uORFs Are Characterized by a Low TE in *T. cruzi*

An open reading frame with repressive potential in the 5′UTR region of a gene could decrease its translation efficiency. To independently assess the repressive effect of the two defined

TABLE 1 | Number of genes presenting uORFs in *T. cruzi* epimastigotes and metacyclic trypomastigote stages.

T. cruzi life cycle stage	Repressive uORFs			Non-repressive	Total analyzed
	uORFo	uORFno	Both		
Epimastigote	111	160	297	3375	6744
Trypomastigote	204	167	231	3333	6750

The number of repressive (either overlapping, non-overlapping or both) and non-repressive uORFS are indicated. Also, the total number of transcripts analyzed for each T. cruzi stage is shown.

categories, the TE values of the genes whose 5′UTR regions contain only repressive uORFno or only uORFo were calculated using Ribodiff and available ribosome profiling data obtained by our group (Smircich et al., 2015). TE values were also obtained for the non-repressive category and for all the genes. Translation efficiency values were then correlated to the presence of repressive uORFs in the 5′UTRs. The results show that the TE is indeed significantly correlated with the presence of uORFs that have repressive potential and with the subcategory to which they belong (**Figure 1A**).

Interestingly, genes containing only uORFo have the lowest TE. As ribosomes translating an uORFo are out of frame, they will read through the main CDS AUG, thus establishing an important translational control mechanism. The effect of naturally occurring uORFo has been poorly described in the literature, there are only a few previous reports on it (Wethmar, 2014). Recently, Fervers et al. observed that in *T. congolense,* as the distance between the uAUG and the AUG of the main CDS decreases, the translation efficiency does too, reaching the maximum decrease when the overlap occurs (Fervers et al., 2018), suggesting that this effect may be shared among trypanosomatids.

In the case of genes containing only uORFno, the drop in observed TE is lower than for those containing only uORFo. However, a clear repressive effect is seen, particularly when compared with the TE of the transcripts with 5′UTRs categorized as non-repressive. The TE of this later group is the highest in the entire comparison, suggesting that they indeed lack repressive uORFs and highlighting the relevance of these elements in translation control. In turn, the TE of all the 5′UTR regions shows an intermediate value between the repressive and non-repressive categories as expected considering the latter results.

In the metacyclic trypomastigote stage, a similar situation is evident even though the data show a higher dispersion, specially for the uORFno-containing genes (**Figure 1B**). It is worth noting that in the metacyclic trypomastigote stage, genome-wide translational repression has been described (Smircich et al., 2015), implying the existence of other regulatory mechanisms that exert a more significant effect on translation control.

It has been reported that the higher the number of uORFs in the 5′UTR of an mRNA, the higher the decrease in its translation efficiency (Chew et al., 2016). However, no correlation between the number of uORFs present in the 5′UTR region and the translation efficiency of the genes was found in our model (**Supplementary Figure 1**).

Globally, the results presented here are in good agreement with those previously obtained in other trypanosomatids, suggesting that the presence of open reading frames with regulatory potential in the 5′UTR regions is a contributing factor to translational control in these organisms. We also demonstrate that overlapping upstream open reading frames achieve the highest level of repression in *T. cruzi*. Additionally, the features we selected for uORF classification were confirmed to behave as good indicators of the repressive potential of uORFs, both in the case of repressive and non-repressive categories. Finally, we produced a dataset of *T. cruzi* genes where uORFs are likely acting as important regulatory elements.

AUG Is the Main Initiation Codon of the Repressive uORF in *T. cruzi*

Pioneering studies initially proposed that uORFs with non-AUG initiation codons have poor translation efficiency (Clements et al., 1988). Later, several reports claimed that the majority of the translational initiation codons for uORFs were non-canonical (not AUG) (Ingolia et al., 2011; Fritsch et al., 2012; Lee et al., 2012). More recently, through the analysis of Ribo-seq data, it has been suggested that the translation efficiency of non-AUG initiating uORFs is low (Michel et al., 2014).

This led us to ask whether the presence of AUG as a start codon of an uORF can modulate the repressive effect on translation efficiency in *T. cruzi*. To find an answer to this question, we evaluated if uORFno initiating from each of the 61 codons would also cause a general effect on TE of the main CDS, as observed for AUG. Putative repressive uORFs were defined by maintaining the requirements to classify uORFs in this category (length, distance to the AUG, distance to the 5′ end) and only changing the criteria for the identity of the initiation codon. Thus, for each codon a set of uORFno was determined while preserving the rest of the requirements as before. Then, each new set of uORFs was correlated with the TE distribution of the genes, as done for before AUG initiating uORFs.

In the epimastigote stage, the translation efficiency of genes with uORFno using AUG as initiation codon is significantly lower (non-parametric Mann–Whitney U test < 0.01) compared to any other codon (**Figure 2A** and **Supplementary Figure 2**). Our results show no evidence of any particular behavior of mRNAs containing uORFs beginning with near cognate codons (NCC) known to be capable of translation initiation in other systems (CUG, UUG, GUG, ACG, AUA, and AUU) (Peabody, 1989; Ivanov et al., 2011). A similar situation is observed for the metacyclic trypomastigote stage (**Figure 2B**), the uORFno starting with AUG also being the only initiation codon associated with statistical significance to a low TE.

This approach allowed us to determine that AUG is the main initiation codon that generates repressive uORFno in *T. cruzi*. This finding further supports our original criteria. It is worth noting that this analysis does not eliminate the possibility that NCC initiate translation in specific uORFs with important regulatory consequences for the affected CDS. Indeed, experimentally assessing the translation of uORFs initiated by NCC would provide interesting insights in this regard. Even

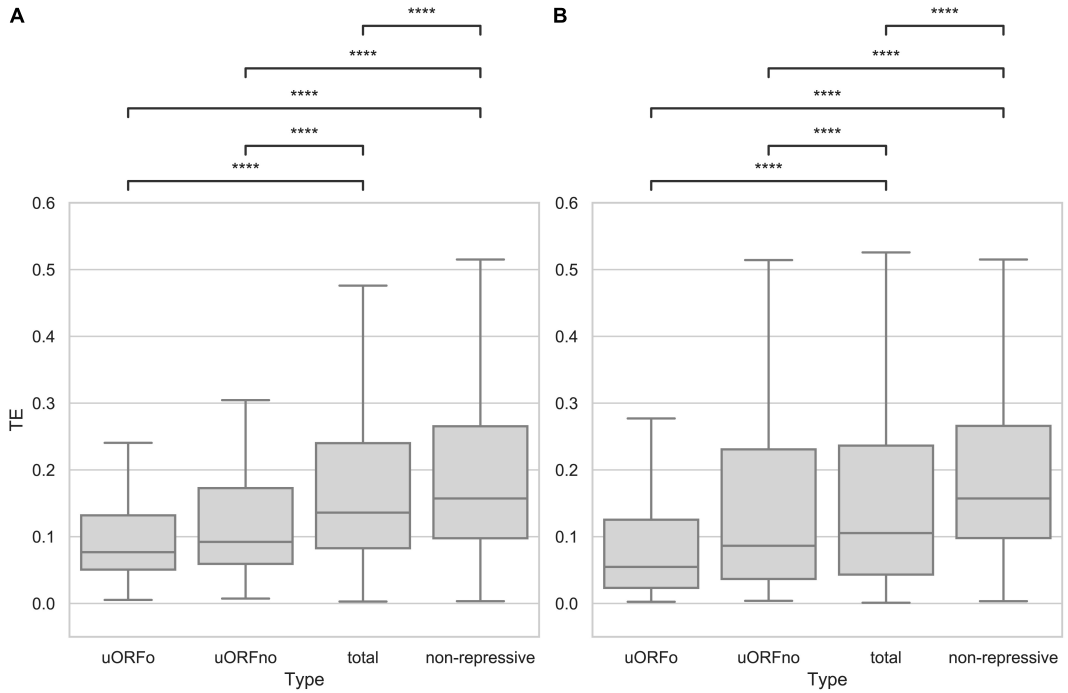

FIGURE 1 | Translation efficiency of genes containing uORFs in *T. cruzi*. Box plots showing the distribution of TE values for genes with 5′UTR containing uORFo, uORFno, all genes and genes presenting UTRs classified as non-repressive. Statistically significant comparisons are indicated (Mann–Whitney *U* $p < 1e\text{-}4$ ****, $p < 1e\text{-}3$ ***, $p < 1e\text{-}2$ **, $p < 0.05$ *). **(A)** epimastigotes. **(B)** metacyclic trypomastigotes.

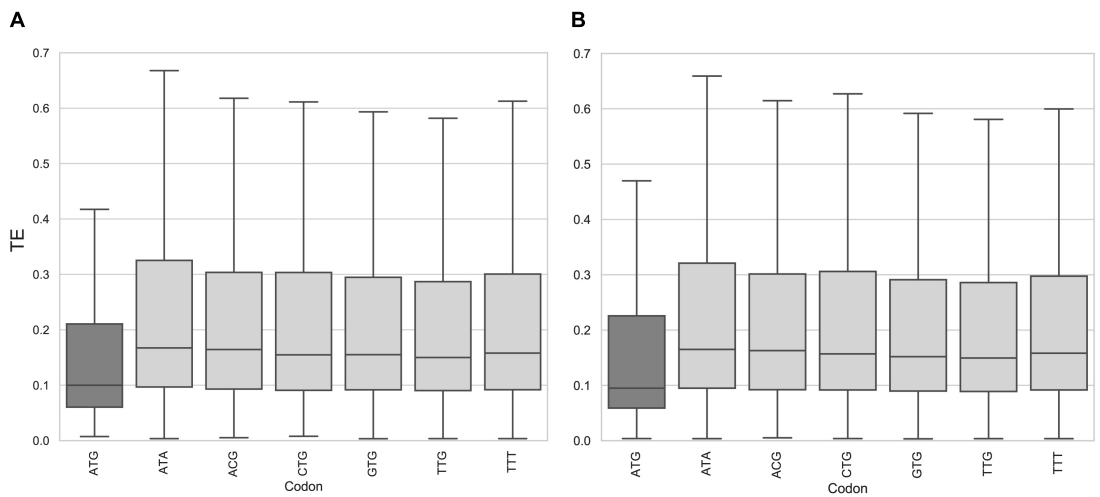

FIGURE 2 | Translation efficiency of genes containing uORFno in *T. cruzi*. Box plots showing the distribution of translation efficiency values for genes with 5′UTR containing uORFno starting at each indicated codon. For each boxplot the procedure to identify uORFno with repressive potential was identical. TTT is used as a control codon (not described to efficiently initiate translation). **(A)** epimastigotes. **(B)** metacyclic trypomastigotes.

though Ribo-seq data has been used to this end in other models, the sensitivity reached by our data for the 5′UTR regions did not allow us to address this issue.

Kozak proposed that translation efficiency is strongly determined by the context of the initiating AUG (Kozak, 1978, 2002). In many organisms including *T. brucei* and *L. major*, only some of the characteristics defined by Kozak for mammalian cells are preserved, mainly the presence of A in the -3 position (Nakagawa et al., 2008). To further address this question in *T. cruzi*, we studied the primary sequence context of the initiator AUG of all main CDSs and repressive uORFno in the epimastigote stage (10 nt flanking the A in position +1 of the CDS). While the overrepresentation of A at position -3 is not as evident as for the other trypanosomatids, a clear purine

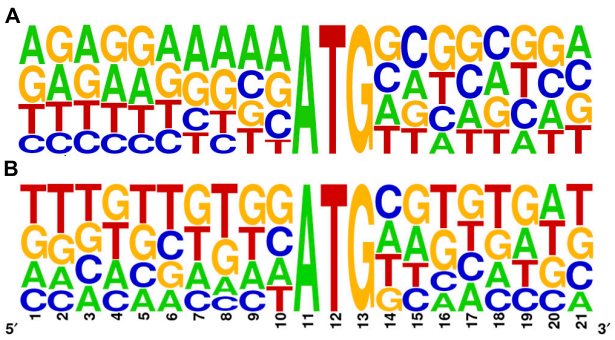

FIGURE 3 | Sequence logo of the AUG initiator sequence in *T. cruzi*. 10 base pairs surrounding the A in position +1 are shown for the main CDS **(A)** and for the longest uORFno **(B)**. Context logo were generated with WebLogo (Crooks et al., 2004).

enrichment is found for the main CDS AUG. However, this is not observed for the uORFs AUG (**Figure 3**), suggesting that initiation driven from the uAUG is not as efficient as from the main CDS AUG.

The Presence of uORFno Is Correlated to the 5′UTR Length

As the presence of a uORFno will set a minimum length for the 5′UTR, we explored the size of this region. We found that the 5′UTR size of the uORFno associated genes is significantly larger (Mann–Whitney *U* test, *p*-value < 0.001) than the rest of the groups, for both the epimastigotes or metacyclic trypomastigote stages, as can be seen in **Figure 4**. Interestingly, the median size of repressive UTRs containing uORFno almost triples the length of the average *T. cruzi* 5′UTR. This finding implies that either the maintenance of uORFno increases the size of these regions, or that the length of the 5′UTR by itself is a determinant of the repressive potential. The latter is not a likely explanation as there is no general genome wide correlation between 5′UTR length and TE (**Supplementary Figure 3**).

In addition, this observation suggests that the minimum length of the 5′UTRs of genes regulated by uORFno might be evolutionary restricted to allow for the presence of the regulatory element. In support of this hypothesis, regions containing uORFo (elements that repress TE through a size independent mechanism), have the shortest 5′UTRs of all the groups.

Genes Associated to Repressive 5′UTRs Have Low Expression in *T. cruzi*

In order to study the association between the presence of repressive uORFs and the translational level of the CDS, we comparatively analyzed the translation rates of genes containing repressive and non-repressive uORFs using our Ribo-Seq data. We found that genes with repressive uORFs have a decreased the number of ribosomal footprints compared to genes with non-repressive UTRs. Accordingly, the latter

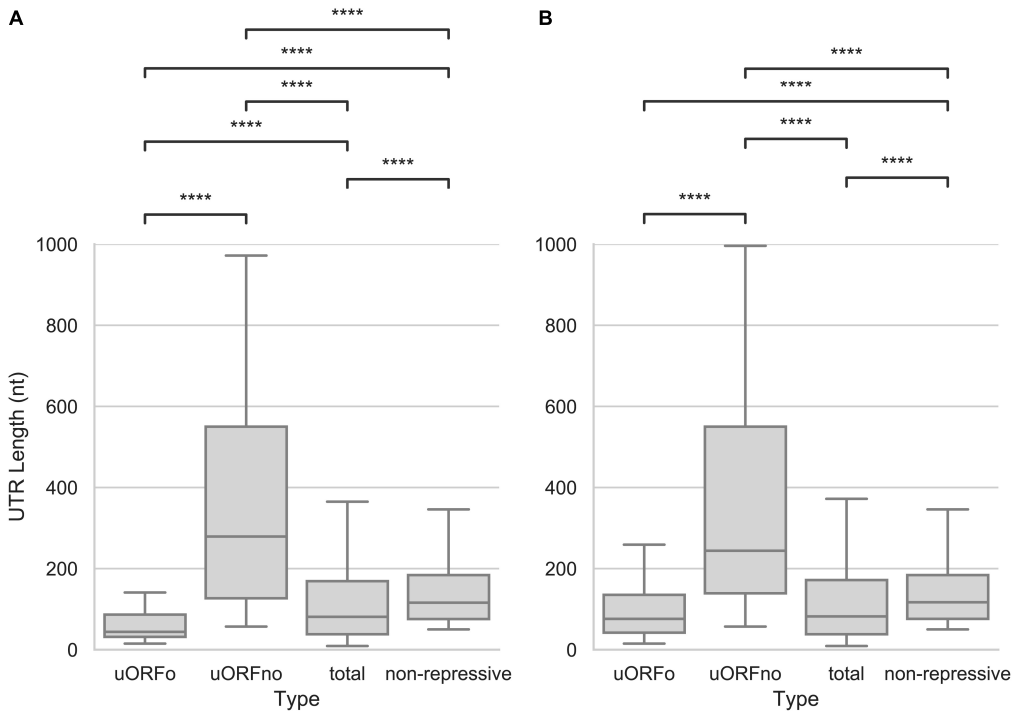

FIGURE 4 | Length of the 5′UTRs in *T. cruzi*. Box plots showing the distribution of 5′UTR lengths for genes with 5′UTR containing uORFo, uORFno, all genes and genes presenting UTRs classified as non-repressive. Statistically significant comparisons are indicated (Mann–Whitney *U* $p < 1e\text{-}4$ ****, $p < 1e\text{-}3$ ***, $p < 1e\text{-}2$ **, $p < 0.05$ *). **(A)** epimastigotes. **(B)** metacyclic trypomastigotes.

are also have the greatest translation rates (**Figure 5A** and **Supplementary Figure 4A**).

This effect is also reflected in the available proteomic data (de Godoy et al., 2012) where most proteins translated from genes containing repressive uORFs are not detected. Indeed, we found a significant underrepresentation of these genes in proteomic databases and an overrepresentation of proteins whose mRNAs do not contain repressive uORFs (Thermo Fisher Scientific test < 0.05, **Figure 5B** and **Supplementary Figure 4B**). As detection in proteomic studies is biased by the relative amount of protein in the sample, these data suggest low levels for proteins coded by repressive uORF containing mRNAs.

Furthermore, genes with non-repressive 5′UTRs show ontology term enrichment with a trend for housekeeping functions such as catabolic processes, cell movement and transport, which are generally associated with high expression protein levels. Conversely, genes with 5′UTR containing repressive uORFs show enrichment in translational elongation

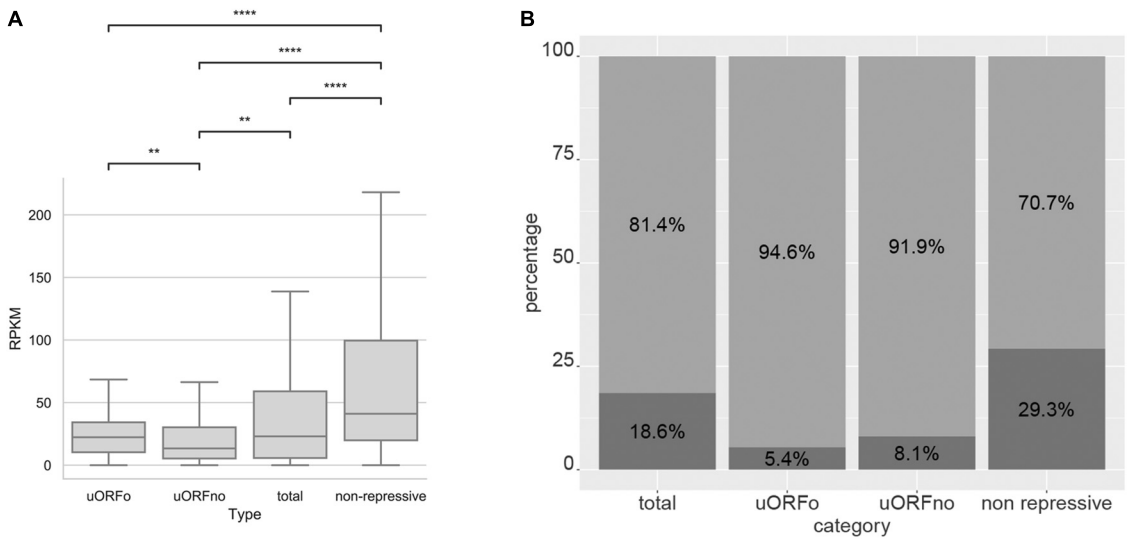

FIGURE 5 | Analysis of the translation of genes with uORFs in *T. cruzi* epimastigotes. **(A)** Ribosomal footprints obtained from Smircich et al. (2015) were analyzed for genes with 5′UTR containing uORFo, uORFno, all genes and genes presenting UTRs classified as non-repressive. Boxplot of ribosomal footprints (RPKM) for genes in each category. Statistically significant comparisons are indicated (Mann–Whitney *U* $p < 1e\text{-}4$ ****, $p < 1e\text{-}3$ ***, $p < 1e\text{-}2$ **, $p < 0.05$ *). **(B)** Fraction of the genes belonging to each category present in the proteomic data of de Godoy et al. (2012) are represented in dark gray. A Fisher's exact test was used to assess over or down representation of the number of detected proteins in proteomic experiments for each category compared to the number detected in the total proteome (all comparisons are significant $p < 0.05$).

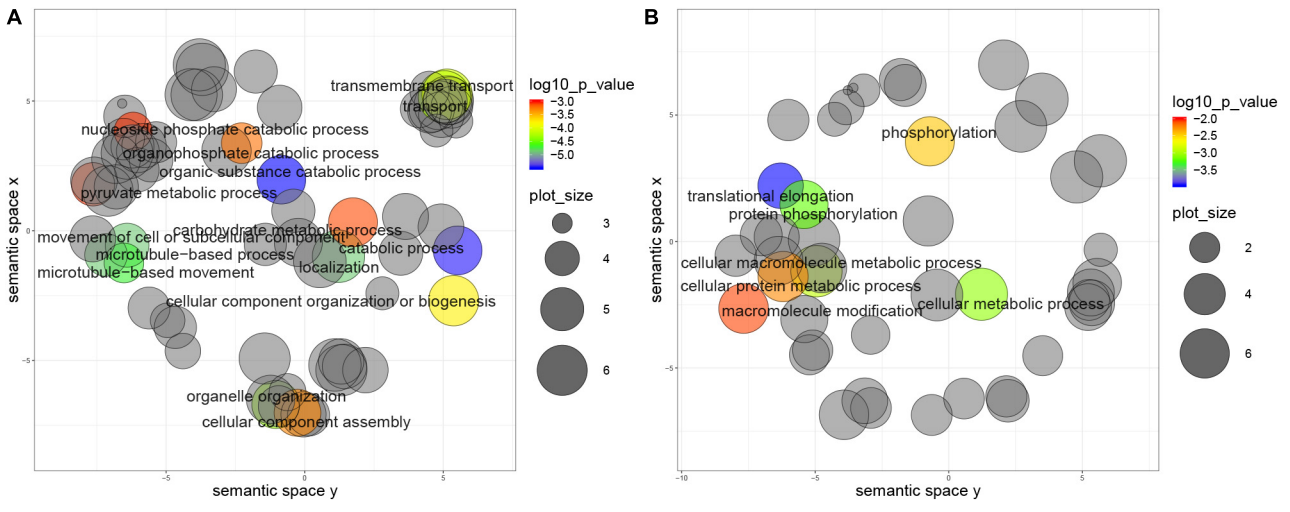

FIGURE 6 | Biological process categories overrepresented in genes with uORFs in *T. cruzi* epimastigotes. Grouping reveals functionally related categories, size shows the frequency of the UniProtKB database category and color indicates a *p* value < 0.01. **(A)** Genes with non-repressive 5′UTRs, **(B)** Genes with repressive 5′UTRs.

and phosphorylation terms (**Figures 6A,B**). Similar results were obtained for the metacyclic trypomastigote stage, so we can conclude that this is not a stage specific characteristic (data not shown).

Finally, to evaluate if alternative 5′UTR processing could provide a mechanism to regulate the presence of uORFs (and thus the TE) between *T. cruzi* life stages, we searched for differential 5′UTR processing between epimatigotes and metacyclic trypomastigotes. First, we observed that most genes either share their main 5′UTR splice site (83% of the 3245 analyzed), or have sites are less than 10 nt apart (60% of the non-identical ones) (**Supplementary Figure 5**). This indicates that the presence or absence of uORFs is not a general mechanism to regulate differential TE between these life-cycle stages. We cannot discard that differential translation efficiency of the uORFs themselves might provide a regulatory mechanism of the main CDS. This requires further investigation as it cannot be assessed with our current data.

FINAL REMARKS

Overall, the results here presented allow us to conclude that uORFs are a frequent mechanism to fine tune translation efficiency in *T. cruzi*. Characteristics intrinsic to the uORF – such as its position within the 5′UTR, size and start codon – influence their ability to exert this effect. Genes with repressive uORFs have low levels of ribosomal footprints and are underrepresented in proteomic data, while the opposite is observed for genes with 5′UTR regions classified as non-repressive. Finally, differential 5′UTR processing does not seem to be a general mechanism to regulate their TE between the parasite life-cycle stages analyzed.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

SR and PS designed the methodology. SR performed the analysis. JS-S, BG, and PS acquired the financial support. PS coordinated the project. All authors wrote and reviewed the manuscript.

FUNDING

This project was supported by the ANII, FCE_3_2016_1_126317; CSIC, I+D research groups program 108725. SR received scholarships from the ANII. SR and PS received financial support from the PEDECIBA.

ACKNOWLEDGMENTS

The authors would like to acknowledge Dr. Martin Ciganda for careful revision of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00166/full#supplementary-material

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Calvo, S. E., Pagliarini, D. J., and Mootha, V. K. (2009). Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc. Natl. Acad. Sci. U.S.A.* 106, 7507–7512. doi: 10.1073/pnas.0810916106

Chen, Y. J., Tan, B. C., Cheng, Y. Y., Chen, J. S., and Lee, S. C. (2010). Differential regulation of CHOP translation by phosphorylated eIF4E under stress conditions. *Nucleic Acids Res.* 38, 764–777. doi: 10.1093/nar/gkp1034

Chew, G. L., Pauli, A., and Schier, A. F. (2016). Conservation of uORF repressiveness and sequence features in mouse, human and zebrafish. *Nat. Commun.* 7:11663. doi: 10.1038/ncomms11663

Clements, J. M., Laz, T. M., and Sherman, F. (1988). Efficiency of translation initiation by non-AUG codons in Saccharomyces cerevisiae. *Mol. Cell Biol.* 8, 4533–4536. doi: 10.1128/mcb.8.10.4533

Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190. doi: 10.1101/gr.849004

de Godoy, L. M., Marchini, F. K., Pavoni, D. P., Rampazzo Rde, C., Probst, C. M., Goldenberg, S., et al. (2012). Quantitative proteomics of Trypanosoma cruzi during metacyclogenesis. *Proteomics* 12, 2694–2703. doi: 10.1002/pmic.201200078

Fervers, P., Fervers, F., Makalowski, W., and Jakalski, M. (2018). Life cycle adapted upstream open reading frames (uORFs) in *Trypanosoma congolense*: a post-transcriptional approach to accurate gene regulation. *PLoS One* 13:e0201461. doi: 10.1371/journal.pone.0201461

Fritsch, C., Herrmann, A., Nothnagel, M., Szafranski, K., Huse, K., Schumann, F., et al. (2012). Genome-wide search for novel human uORFs and N-terminal protein extensions using ribosomal footprinting. *Genome Res.* 22, 2208–2218. doi: 10.1101/gr.139568.112

Griffin, E., Re, A., Hamel, N., Fu, C., Bush, H., McCaffrey, T., et al. (2001). A link between diabetes and atherosclerosis: glucose regulates expression of CD36 at the level of translation. *Nat. Med.* 7, 840–846. doi: 10.1038/89969

Ingolia, N. T., Lareau, L. F., and Weissman, J. S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* 147, 789–802. doi: 10.1016/j.cell.2011.10.002

Ivanov, I. P., Firth, A. E., Michel, A. M., Atkins, J. F., and Baranov, P. V. (2011). Identification of evolutionarily conserved non-AUG-initiated N-terminal extensions in human coding sequences. *Nucleic Acids Res.* 39, 4220–4234. doi: 10.1093/nar/gkr007

Jaeger, L. H., and Brandao, A. (2011). The composition of upstream open reading frames (uORF) in four genes from Trypanosoma cruzi typical strains. *Parasitol. Res.* 109, 1205–1208. doi: 10.1007/s00436-011-2430-4

Jensen, B. C., Ramasamy, G., Vasconcelos, E. J., Ingolia, N. T., Myler, P. J., and Parsons, M. (2014). Extensive stage-regulation of translation revealed by ribosome profiling of *Trypanosoma brucei*. *BMC Genomics* 15:911. doi: 10.1186/1471-2164-15-911

Johnson, N. R., Yeoh, J. M., Coruh, C., and Axtell, M. J. (2016). Improved placement of multi-mapping small RNAs. *G3* 6, 2103–2111. doi: 10.1534/g3.116.030452

Kozak, M. (1978). How do eucaryotic ribosomes select initiation regions in messenger RNA? *Cell* 15, 1109–1123. doi: 10.1016/0092-8674(78)90039-9

Kozak, M. (1987). Effects of intercistronic length on the efficiency of reinitiation by eucaryotic ribosomes. *Mol. Cell Biol.* 7, 3438–3445. doi: 10.1128/mcb.7.10.3438

Kozak, M. (2001). Constraints on reinitiation of translation in mammals. *Nucleic Acids Res.* 29, 5226–5232. doi: 10.1093/nar/29.24.5226

Kozak, M. (2002). Pushing the limits of the scanning mechanism for initiation of translation. *Gene* 299, 1–34. doi: 10.1016/s0378-1119(02)01056-9

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10:R25. doi: 10.1186/gb-2009-10-3-r25

Lawless, C., Pearson, R. D., Selley, J. N., Smirnova, J. B., Grant, C. M., Ashe, M. P., et al. (2009). Upstream sequence elements direct post-transcriptional regulation of gene expression under stress conditions in yeast. *BMC Genomics* 10:7. doi: 10.1186/1471-2164-10-7

Lee, S., Liu, B., Lee, S., Huang, S. X., Shen, B., and Qian, S. B. (2012). Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. *Proc. Natl. Acad. Sci. U.S.A.* 109, E2424–E2432. doi: 10.1073/pnas.1207846109

Li, Y., Shah-Simpson, S., Okrah, K., Belew, A. T., Choi, J., Caradonna, K. L., et al. (2016). Transcriptome remodeling in *Trypanosoma cruzi* and human cells during intracellular infection. *PLoS Pathog.* 12:e1005511. doi: 10.1371/journal.ppat.1005511

Liao, Y., Smyth, G. K., and Shi, W. (2013). The *Subread aligner*: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* 41:e108. doi: 10.1093/nar/gkt214

Lidani, K. C. F., Andrade, F. A., Bavia, L., Damasceno, F. S., Beltrame, M. H., Messias-Reason, I. J., et al. (2019). Chagas disease: from discovery to a worldwide health problem. *Front. Public Health* 7:166. doi: 10.3389/fpubh.2019.00166

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* 17, 10–12.

McCarthy, J. E. (1998). Posttranscriptional control of gene expression in yeast. *Microbiol. Mol. Biol. Rev.* 62, 1492–1553. doi: 10.1128/mmbr.62.4.1492-1553.1998

Michel, A. M., Andreev, D. E., and Baranov, P. V. (2014). Computational approach for calculating the probability of eukaryotic translation initiation from ribo-seq data that takes into account leaky scanning. *BMC Bioinformatics* 15:380. doi: 10.1186/s12859-014-0380-4

Nakagawa, S., Niimura, Y., Gojobori, T., Tanaka, H., and Miura, K. (2008). Diversity of preferred nucleotide sequences around the translation initiation codon in eukaryote genomes. *Nucleic Acids Res.* 36, 861–871. doi: 10.1093/nar/gkm1102

Peabody, D. S. (1989). Translation initiation at non-AUG triplets in mammalian cells. *J. Biol. Chem.* 264, 5031–5035.

R Core Team, (2013). *R: A Language and Environment for Statistical Computing*. Vienna: R Core Team.

Radio, S., Fort, R. S., Garat, B., Sotelo-Silveira, J., and Smircich, P. (2018). UTRme: a scoring-based tool to annotate untranslated regions in trypanosomatid genomes. *Front. Genet.* 9:671. doi: 10.3389/fgene.2018.00671

Rajkowitsch, L., Vilela, C., Berthelot, K., Ramirez, C. V., and McCarthy, J. E. (2004). Reinitiation and recycling are distinct processes occurring downstream of translation termination in yeast. *J. Mol. Biol.* 335, 71–85. doi: 10.1016/j.jmb.2003.10.049

Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: the european molecular biology open software suite. *Trends Genet.* 16, 276–277. doi: 10.1016/s0168-9525(00)02024-2

Siegel, T. N., Tan, K. S., and Cross, G. A. (2005). Systematic study of sequence motifs for RNA trans splicing in Trypanosoma brucei. *Mol. Cell Biol.* 25, 9586–9594. doi: 10.1128/MCB.25.21.9586-9594.2005

Smircich, P., Eastman, G., Bispo, S., Duhagon, M. A., Guerra-Slompo, E. P., Garat, B., et al. (2015). Ribosome profiling reveals translation control as a key mechanism generating differential gene expression in Trypanosoma cruzi. *BMC Genomics* 16:443. doi: 10.1186/s12864-015-1563-8

Soldatos, T. G., O'Donoghue, S. I., Satagopam, V. P., Jensen, L. J., Brown, N. P., Barbosa-Silva, A., et al. (2010). Martini: using literature keywords to compare gene sets. *Nucleic Acids Res.* 38, 26–38. doi: 10.1093/nar/gkp876

Supek, F., Bosnjak, M., Skunca, N., and Smuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* 6:e21800. doi: 10.1371/journal.pone.0021800

Tyler, K. M., and Engman, D. M. (2001). The life cycle of Trypanosoma cruzi revisited. *Int. J. Parasitol.* 31, 472–481. doi: 10.1016/s0020-7519(01)00153-9

Vasquez, J. J., Hon, C. C., Vanselow, J. T., Schlosser, A., and Siegel, T. N. (2014). Comparative ribosome profiling reveals extensive translational complexity in different *Trypanosoma brucei* life cycle stages. *Nucleic Acids Res.* 42, 3623–3637. doi: 10.1093/nar/gkt1386

Vattem, K. M., and Wek, R. C. (2004). Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 101, 11269–11274. doi: 10.1073/pnas.0400541101

Vilela, C., and McCarthy, J. E. (2003). Regulation of fungal gene expression via short open reading frames in the mRNA 5′untranslated region. *Mol. Microbiol.* 49, 859–867. doi: 10.1046/j.1365-2958.2003.03622.x

Wethmar, K. (2014). The regulatory potential of upstream open reading frames in eukaryotic gene expression. *Wiley Interdiscip. Rev. RNA* 5, 765–778. doi: 10.1002/wrna.1245

Zhong, Y., Karaletsos, T., Drewe, P., Sreedharan, V. T., Kuo, D., Singh, K., et al. (2017). RiboDiff: detecting changes of mRNA translation efficiency from ribosome footprints. *Bioinformatics* 33, 139–141. doi: 10.1093/bioinformatics/btw585

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

ORIGINAL RESEARCH
published: 05 March 2020
doi: 10.3389/fgene.2020.00144

Valine-Induced Isoleucine Starvation in *Escherichia coli* K-12 Studied by Spike-In Normalized RNA Sequencing

Bertil Gummesson[1], Shiraz Ali Shah[1†], Alexander Skov Borum[1], Mathias Fessler[1†], Namiko Mitarai[2], Michael Askvad Sørensen[1] and Sine Lo Svenningsen[1*]*

[1] Department of Biology, University of Copenhagen, Copenhagen, Denmark, [2] Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark

OPEN ACCESS

Edited by:
Sanjeev Kumar Srivastava, Mitchell Cancer Institute, United States

Reviewed by:
Jue Wang, University of Wisconsin-Madison, United States
Michael Ibba, The Ohio State University, United States

***Correspondence:**
Michael Askvad Sørensen
MAS@bio.ku.dk
Sine Lo Svenningsen
SLS@bio.ku.dk

†Present address:
Shiraz Ali Shah, Copenhagen Prospective Studies on Asthma in Childhood, Copenhagen University Hospital, Herlev-Gentofte, Gentofte, Denmark
Mathias Fessler, DTU Environment, Technical University of Denmark, Kongens Lyngby, Denmark

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *06 September 2019*
Accepted: *07 February 2020*
Published: *05 March 2020*

Citation:
Gummesson B, Shah SA, Borum AS, Fessler M, Mitarai N, Sørensen MA and Svenningsen SL (2020) Valine-Induced Isoleucine Starvation in Escherichia coli K-12 Studied by Spike-In Normalized RNA Sequencing. Front. Genet. 11:144. doi: 10.3389/fgene.2020.00144

Escherichia coli cells respond to a period of famine by globally reorganizing their gene expression. The changes are known as the stringent response, which is orchestrated by the alarmone ppGpp that binds directly to RNA polymerase. The resulting changes in gene expression are particularly well studied in the case of amino acid starvation. We used deep RNA sequencing in combination with spike-in cells to measure global changes in the transcriptome after valine-induced isoleucine starvation of a standard *E. coli* K12 strain. Owing to the whole-cell spike-in method that eliminates variations in RNA extraction efficiency between samples, we show that ribosomal RNA levels are reduced during isoleucine starvation and we quantify how the change in cellular RNA content affects estimates of gene regulation. Specifically, we show that standard data normalization relying on sample sequencing depth underestimates the number of down-regulated genes in the stringent response and overestimates the number of up-regulated genes by approximately 40%. The whole-cell spike-in method also made it possible to quantify how rapidly the pool of total messenger RNA (mRNA) decreases upon amino acid starvation. A principal component analysis showed that the first two components together described 69% of the variability of the data, underlining that large and highly coordinated regulons are at play in the stringent response. The induction of starvation by sudden addition of high valine concentrations provoked prominent regulatory responses outside of the expected ppGpp, RpoS, and Lrp regulons. This underlines the notion that with the high resolution possible in deep RNA sequencing analysis, any different starvation method (e.g., nitrogen-deprivation, removal of an amino acid from an auxotroph strain, or valine addition to *E. coli* K12 strains) will produce measurable variations in the stress response produced by the cells to cope with the specific treatment.

Keywords: stringent response, deep RNA sequencing, whole-cell spike-in normalization, ribosomal RNA degradation, transcriptome, ppGpp, gene expression

INTRODUCTION

During stress conditions, cells of *Escherichia coli* (*E. coli*) impose dramatic changes in their transcriptional profile and proteome to combat stressors. The cells ensure that genes important to overcome the stress are turned on and other redundant and energy-demanding gene products, such as genes of the protein synthesis machinery [*i.e.,* those encoding the ribosomes, transfer RNAs (tRNAs) and factors required for translation] are down-regulated. The rapid re-orchestration of the transcriptome in *E. coli* occurs on the timescale of a few minutes, and is aided by the small molecules guanosine tetra- and pentaphosphate, herein collectively referred to as ppGpp. This physiological response is called the stringent response (Ryals et al., 1982; Cashel et al., 1996) and has become a model system for studies of bacterial stress responses. Together with the protein DksA, ppGpp binds two sites on RNA polymerase, which affects promoter selectivity and reduces the ribosomal RNA (rRNA) promoter clearing rates (Artsimovitch et al., 2004; Gummesson et al., 2013; Ross et al., 2016). The nucleotide ppGpp is produced when amino acids become limiting and upon starvation for many different kinds of nutrients as well as by other circumstances restricting growth (Cashel et al., 1996). In *E. coli*, the synthesis of ppGpp is mediated by two related proteins, RelA and SpoT; each requiring different signals for activation. The RelA protein is associated with uncharged tRNA and the synthesis of ppGpp is triggered when the translating ribosome binds a RelA-tRNA complex at the starving A-site codon (Haseltine and Block, 1973; Winther et al., 2018). The SpoT protein is bi-functional; besides synthesizing ppGpp, SpoT can hydrolyse ppGpp to guanosine diphosphate and pyrophosphate (Murray and Bremer, 1996), thus allowing a way out of stringency when conditions allow.

The global effect of ppGpp on transcription has previously been studied upon starvation for the amino acid serine (Durfee et al., 2008) or isoleucine (Traxler et al., 2008; Traxler et al., 2011). These studies have in common that they utilized the well-established expression microarrays as their read out, the best technology available for genome-wide analysis at the time. However, the much more sensitive technique of deep RNA sequencing (RNAseq) has emerged as a standard method to measure globally the relative abundance of RNA species in the cell, and offers a superior dynamic range for measuring variability in the levels of expressed transcripts (Wang et al., 2009; Croucher and Thomson, 2010). The effects of ppGpp on global transcriptional regulation without concomitant starvation has recently been studied using RNAseq, and resulted in a substantial expansion of the genes that can be assigned to the ppGpp-controlled regulon (Sanchez-Vazquez et al, 2019).

The long-lived house-keeping RNAs, rRNA, and tRNA, constitute the vast majority of the total RNA in the cell (>95%) (Bremer and Dennis, 1996). For this reason, rRNA and sometimes tRNA are generally removed prior to RNAseq, or not included on microarrays, unless they are the specific focus of the study. One goal of our work was to obtain data on the response to amino acid starvation in *E. coli* that includes the changes in the whole transcriptome, including the most abundant RNAs, and to analyze how inclusion of all RNA may enhance the current understanding of the well-studied stringent response. In connection with this goal came the need to quantify transcripts without making assumptions about the total RNA content of the cells before and after starvation. Typical transcriptome analyses, whether done by microarray or RNAseq, rely on the assumption that the total amount of RNA is constant across different sample conditions. However, while rRNA and tRNAs are generally believed to be stable during exponential growth (Baracchini and Bremer, 1987), the familiar way of thinking of these RNAs as stable in an absolute sense has been questioned for some time (Deutscher, 2003). Our previous work shows that a substantial fraction of the tRNA and rRNA pools in the cell is rapidly degraded upon amino acid starvation (Svenningsen et al., 2017; Fessler et al., 2020), suggesting that the total RNA content of *E. coli* cells may decrease appreciably under this condition. Given the global changes in gene expression and the possibility that total RNA levels may decrease upon amino acid starvation we reasoned that a normalization method that is independent of any assumptions about cellular RNA content would be important for accurate detection of gene expression changes during the stringent response. Therefore, we chose to normalize the sample sequencing reads using a spike-in culture for reference. Spike-in, in the form of *in vitro* synthesized RNA, has been used in many experiments for normalization of transcriptional activity (see e.g., Schena et al., 1995; Bartholomäus et al., 2016; Gorochowski et al., 2019) and to verify the accuracy of RNA preparation protocols (see e.g., Jones et al., 2015; Ju et al., 2019). However, *in-vitro*-transcribed spike-in RNA is added after the extraction of the experimental RNA and quantification of transcription rates assume an equally efficient extraction of RNA from each sample (Gorochowski et al., 2019). Or, if the spike-in transcripts are added per mass of RNA in each sample, the underlying assumption is that cells contain equal amounts of RNA at the different conditions. The whole-cell spike-in approach is often used in microbiome studies to quantify cell numbers (Hornung et al., 2019) but has not, to our knowledge, been used outside our research group for quantification of RNA (Svenningsen et al., 2017). The benefit of the whole-cell spike-in approach we use here is that it allows normalization directly to the concentration of bacteria in each sample [as measured by optical density (OD)], without making any assumptions about the RNA content of the cells. For comparison, we also normalized our data set using the conventional approach of normalizing the data based on the sequencing depth obtained for each sample. The analysis of the transcriptome of isoleucine-starved cells normalized by the two methods reveal that the regulon responding negatively to starvation is much larger than what is detected using a conventionally normalized RNAseq transcriptome, and the regulon responding positively is correspondingly smaller. This observation relates to a greater turnover of total RNA in starved cells than previously anticipated, and the spike-in approach enabled us to quantify the loss of rRNAs and total messenger RNA (mRNA) during starvation relative to the levels during steady-state growth. Furthermore, principal component analysis

of the stringent response transcriptome reveals two predominant temporal gene profiles that are enriched for classes of genes with related biological functions. Finally, it was evident that isoleucine starvation induced by L-valine has transcriptional consequences that are separate from the general stringent response of amino-acid-starved cells controlled by ppGpp.

MATERIALS AND METHODS

Strains, Media, and Growth Condition

The wild-type strain *E. coli* K-12 MAS1081 (MG1655 rph^+ $gatC^+$ $glpR^+$) were grown in flasks at 37°C at 200 rpm in morpholinepropanesulfonic acid (MOPS) minimal medium (Neidhardt et al., 1974) supplemented with 0.2% glucose. Cell growth was monitored spectrophotometrically by optical density at 436 nm (OD_{436}) and cultures were grown for at least nine generations in exponential phase before sampling. Isoleucine starvation was induced by adding L-valine to a final concentration of 400 μg/ml (Leavitt and Umbarger, 1962). The small RNA (sRNA) *qrr2* from *Vibrio cholerae* was cloned downstream of the T7 promoter in the vector pET11a (XbaI/Bpu1102I) and transformed into *E. coli* BL21 (DE3) to yield the spike-in strain BKG3; 100 μg/ml ampicillin was used to maintain the plasmid and the expression of Qrr2 was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Rifampicin was added to a final concentration of 300 μg/ml immediately after the last isoleucine starvation sample to block transcription initiation.

Spike-In and RNA Extraction

To preserve cellular RNA, bacterial culture samples were harvested by mixing with 1/6 vol of a stop-solution composed of 5% water-saturated phenol in ethanol at 0°C (Bernstein et al., 2002). All samples were kept at 0°C until the final sample had been harvested. Prior to total RNA extraction, a volume of spike-in culture corresponding to 1% of the experimental culture was added to each sample, based on sample OD. The volume of spike-in cell culture used was calculated according to Equation 1 (as described in Stenum et al., 2017).

$$V_{spike-in} = \frac{0.01 * V_{sample} * OD_{sample}}{OD_{spike-in}} \quad (1)$$

RNA was isolated using a hot phenol extraction method. Briefly, cell pellets were mixed with resuspension solution (0.3 M sucrose, 0.01 M sodium acetate pH 4.5, 0°C), then with lysis solution [2% sodium dodecyl sulfate (SDS), 0.01 M sodium acetate pH 4.5] and finally with hot acidic phenol [pH 4.3, 65° C (Fisher BioReagents)]. The mixture was snap-freezed in liquid nitrogen and centrifuged, and the aqueous phase was re-extracted by phenol (65°C) and frozen in liquid nitrogen one more time. RNA was precipitated with 2.5 vol ethanol and 0.1 vol sodium acetate (3M, pH 4.7) at −80°C overnight. Precipitated RNA was pelleted, washed with 70% ethanol, and re-suspended in nuclease-free H_2O. The remaining DNA was removed by DNaseI treatment (Roche), according to the manufacturers manual. RNA integrity (16S and 23S rRNA) was verified by agarose gel electrophoresis.

Northern Blot

An aliquot of total RNA was mixed with 3 vol loading dye (8 M urea, 6% formaldehyde, bromophenol blue) and fractionated by electrophoresis through a 1% MOPS-buffered agarose gel prepared with 6% formaldehyde. The RNA was transferred from the gel onto a Hybond-N+ membrane by capillary transfer overnight and was fixed to the membrane by 0.12 J/cm^2 of UV light in a Hoefer UVC 500 UV crosslinker. Membranes were pre-hybridized for one hour at 42°C in 6 ml hybridization solution [0.09 M NaCl, 0.05 M NaH_2PO_4 (pH 7.7), 5 mM ethylenediaminetetraacetic acid (EDTA), 5x Denhardt's solution, 0.5% (w/v) SDS, 100 mg/ml sheared, denatured herring sperm DNA]. Hybridization of the immobilized RNA was performed at 42°C overnight in the same solution as above with 40 pmol ^{32}P 5'end-labeled oligo-DNA probe (γ-[^{32}P]-ATP; PerkinElmer). DNA-oligos used were complementary to a sequence in the 5S rRNA, 16S rRNA, 23S rRNA, or Qrr2, probe sequences are listed in **Supplementary Table S11**. Membranes were washed several times in 0.3 M NaCl, 30 mM sodium citrate, 0.1% SDS at room temperature prior to exposure to a phosphor-imaging screen. The radioactivity present in specific bands was measured on a Typhoon phosphor Imager FLA7000 (GE Healthcare) at 100 microns. Membranes were stripped of hybridized probes with 90–95°C stripping buffer (0.1% SDS, 18 mM NaCl, 1 mM NaH_2PO_4, 0.1 mM EDTA) under shaking until no more radioactivity could be detected on the blot by a Geiger-Müller tube. The program ImageQuant TL 8.2 was used to quantify each band on the phosphor-imaging screen. The quantified intensity on each rRNA band were then divided with the values from Qrr2 in the same lane and this ratio is plotted relative to the three samples harvested immediately before inducing starvation.

Quantitative Reverse Transcription PCR

First-strand complementary DNA (cDNA) was reverse transcribed from 1 μg of total RNA with Thermo Scientific RevertAid RT Kit (#K1691) using the supplied random hexamer primers. As control for genomic DNA contamination, a reaction with no reverse transcriptase was included for each sample (RT-). A 1/10,000 to 1/25 fraction of the total synthesized cDNA was combined with SsoAdvanced Universal SYBR Green Supermix (Bio-Rad) and analyzed in triplicate by quantitative reverse transcribed PCR (qRT-PCR) using the QuantStudio 3 system (Applied Biosystems). Thermal cycling conditions used were 95° C for 30 s followed by 40 cycles of 95°C for 15 s, 60°C for 1 min. A final melting-curve cycle was performed to check for amplification artifacts starting at 95° for 15 s, 60° for 1 min, followed by a dissociation step to 95°C with 0.15°C/s increments. The relative levels of RNA is calculated as the signal ratio between the target transcript and one of the reference genes from the spike-in plasmid, namely *bla*, using the formula: $2^{-(\Delta\Delta^{CT})}$ where $\Delta\Delta C_T = (C_{T,target} - C_{T,bla})_{timex} - (C_{T,target} - C_{T,bla})$

time 0 (steady state), as previously described (Livak and Schmittgen, 2001). Primer sequences for target genes and control gene are listed in **Supplementary Table S11**.

RNA Sequencing and Data Analysis

The RNA used for RNAseq was harvested, spiked-in, and extracted as described above; 1–1.5 μg of total RNA from each sample was sent to the GATC BIOTECH facility, European Genome and Diagnostics Centre, Konstanz, Germany for library preparation and RNA sequencing. RNA quality was assessed using an Agilent 2100 Bioanalyzer/Advanced Analytical Technologies Fragment Analyzer. Strand-specific cDNA libraries were prepared according to Illumina's protocols without prior rRNA depletion. RNAseq experiments were performed on an Illumina HiSeq using a paired-end read length of 2x50 bp. Twenty-two to 29 million paired-reads were obtained per sample. GATC BIOTECH initially processed the raw read files, removing adapters prior to delivery. Then the files were uploaded to the Galaxy web platform and we used the public server at usegalaxy.org to analyze the data (Afgan et al., 2018). The files were checked using fastQC[1]. The reads were then aligned to *Escherichia coli* str. K-12 substr. MG1655 (RefSeq NC_U00096.3) using bwa-aln (version 0.7.15.2 with default parameters) (Li and Durbin, 2009). Reads were counted using htseq-count (version 0.6.1p1) (Anders et al., 2015). In parallel, the reads were aligned (bwa-aln) to the reference sequence of the plasmid harboring the spike-in genes and raw read counts mapping to three features (*qrr2, bla,* and antisense-*lacI*), and counted using htseq-count and summed to give the plasmid spike-in reads for a given sample. Raw read counts were then normalized to gene size prior to normalization to spike-in reads to give RPKSP, Reads Per Kilobase of gene per 10 kilobase of spike-in as shown in Equation 2.

$$RPKSP = \frac{(Gene-specific \quad Reads \quad Per \quad Kilobase)}{(Spike-in \quad Reads \quad Per \quad Kilobase/10.000)} \quad (2)$$

We emphasize that the order in which the raw reads were aligned to the *E. coli* chromosome and to the spike-in plasmid did not change the results. Specifically, the same results were obtained when the raw reads were separately aligned to the plasmid and the chromosome as when the alignment was carried out sequentially (i.e., reads were first aligned to the chromosome and remaining reads were aligned to the plasmid).

The raw read counts were also normalized according to the standard method to give RPKM, Reads Per Kilobase Million as shown in Equation 3.

$$RPKM = \frac{(Gene-specific \quad Reads \quad Per \quad Millionreads)}{(Size \quad of \quad Specific \quad gene(kb))} \quad (3)$$

High-throughput sequencing data has been deposited in NCBI's Gene Expression Omnibus (Edgar et al., 2002) and are accessible through GEO Series accession number GSE136753[2].

[1] Andrews, S. FastQC: a quality control tool for high throughput sequence data (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/).

[2] (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE136753)

Transcriptome Data Filtering

We initially applied some filtering of the normalized transcriptomic data (RPKSP and RPKM) in order to quantify the magnitude of fold differences in transcriptional regulation upon starvation for isoleucine. **i**) All genes that were neither sequenced in steady state nor in starvation were filtered out (50 genes). **ii**) Transcripts in the triplicate steady-state samples that either had low average normalized reads or no reads in combination with either low average normalized reads or no reads in the four starvation samples were filtered out (112 genes). These transcripts did not yield any computable fold differences between steady-state growth and starvation. One feature, the gene *lacI*, was present on both the *E. coli* chromosome and the spike-in plasmid. *lacI* was therefore excluded from our analysis. A third filtering step was applied in the comparison of fold differences between RPKM and RPKSP normalization at 10 and 80 min starvation, **iii**) genes where the fold change at both the 10 and 80 min time points relative to the steady-state average could not be calculated due to a lack of coverage were omitted (60 genes). In total 4,048 transcripts were assessed, i.e., 95% of the annotated genes in the *Escherichia coli* str. K-12 substr. MG1655 (RefSeq NC_U00096.3) reference genome. The average standard deviation between the three steady-state measurements of each of the 4,048 transcripts was 25%. The normalized sequencing reads, including omitted genes, are reported in **Supplementary Data Sheet 1** (RPKSP) and **Supplementary Data Sheet 2** (RPKM). For the analysis of the variance among steady-state samples as a function of gene length (**Supplementary Table S2**), we also applied a filter; the analysis was restricted to only consider genes where at least two of the triplicate steady-state samples had detectable transcripts. This yielded 3.979 transcripts for analysis.

Principal Component Analysis

Principal component analysis (PCA) (Abdi and Williams, 2010) was performed on RPKSP-normalized reads of the steady-state samples and the starvation time series in **Supplementary Data Sheet 1**. In order to focus on the temporal profile of expression changes and not the absolute expression level, the number of reads for each data point were normalized to the average number of reads mapped to the corresponding gene for the seven time points. PCA was performed and visualized in MATLAB (MATLAB, Release R2016b). PC1 and PC2 captured 48 and 21% of the variability of the data, respectively. PC3 captured 9% of the variability, but the PC3 vector showed large variability among the three steady-state samples, indicating that it captured a trend that is due to sampling error. Therefore, we focused on the first two principal components. For the enrichment analysis, the Enrichment tool in the SmartTable of the EcoCyc webserver (Karp et al., 2014; Keseler et al., 2017) was used with the options of "Fisher Exact" and "Benjamini-Hochberg Correction" on "Biological Process" gene ontology terms.

Ecocyc Omics Dashboard Tool

Genes in **Supplementary Data Sheet 1** and the $\log_2$-fold induction ratios of the data points in the starvation time series

were imported as a SmartTable in Ecocyc (Karp et al., 2014; Keseler et al., 2017) and analyzed using the Omics Dashboard Tool (Paley et al., 2017). The Dashboard Biosynthesis shown in **Figure 8A** was modified to only show the seven largest sub-systems of biosynthetic genes. In addition, the group of aminoacyl-tRNA synthetases was added manually by curating and extracting the relevant genes from the Biosynthesis sub-system "Others." Genes belonging to the arginine biosynthesis sub-system were exported and their induction ratio at the 5 min time point after starvation are shown alongside data on the same genes extracted from the dataset published by Sanchez-Vazquez and co-workers (Sanchez-Vazquez et al., 2019).

RESULTS

Experimental Approach and Provoking Amino Acid Starvation

To evoke amino acid starvation in cultures of *E. coli* K-12, we grew MAS1081 (MG1655; *rph*$^+$ *gatC*$^+$ *glpR*$^+$) in MOPS-buffered minimal medium supplemented with 0.2% glucose and starved for the amino acid isoleucine by adding excess L-valine. The K-12 strain of *E. coli* harbors a frameshift mutation in *ilvGM*, inactivating one of three isozymes in the valine and isoleucine biosynthetic pathways, while the other two isozymes, *ilvBN* and *ilvIH*, are susceptible to feedback inhibition by L-valine (Valle et al., 2008). High concentrations of L-valine therefore renders *E. coli* K-12 auxotrophic for isoleucine (Leavitt and Umbarger, 1962). Three samples were collected during steady-state growth immediately before starvation and five samples in total were collected in a time series; 5, 10, 20, 40, and 80 min after L-valine addition, resulting in a total of eight samples (**Figure 1A**). A culture of spike-in *E. coli* cells was grown in a parallel, which was not exposed to L-valine (**Figure 1B**). The spike-in cells carry an inducible plasmid and express three transcripts that are not present in the wildtype strain, namely a *V. cholerae* sRNA (*qrr2*), an antibiotic marker (*bla*), and an antisense transcript of *lacI*, from the plasmid. The spike-in cells were induced with IPTG for approximately four generations before they were mixed with the experimental samples in a 1:100 ratio based on OD. We added spike-in cells to the experimental samples prior to total RNA extraction to ensure that variations in RNA recovery, cDNA synthesis, and sequencing depth between the samples would be reflected in the numbers of spike-in reads. By using this approach, we were able to, very precisely, quantify the relative changes in the transcriptome during the experiment, while we lost the information about absolute amounts of transcripts mapped, which is only obtainable by addition of *in vitro* transcribed spike-in RNA after sample preparation (Gorochowski et al., 2019).

Overview of Spike-In Methodology and RNAseq Data

RNAseq libraries were prepared from the eight samples collected during the isoleucine starvation time series. Illumina sequencing results produced 22 to 29 million reads per sample and the proportion of uniquely mapped reads to the *E. coli* genome (RefSeq NC_U00096.3) was at least 97.8% for all samples. The reads from each sample were mapped in parallel to the spike-in plasmid reference sequence (**Supplementary Table S1**). The volume of spike-in cells added to each sample prior to RNA

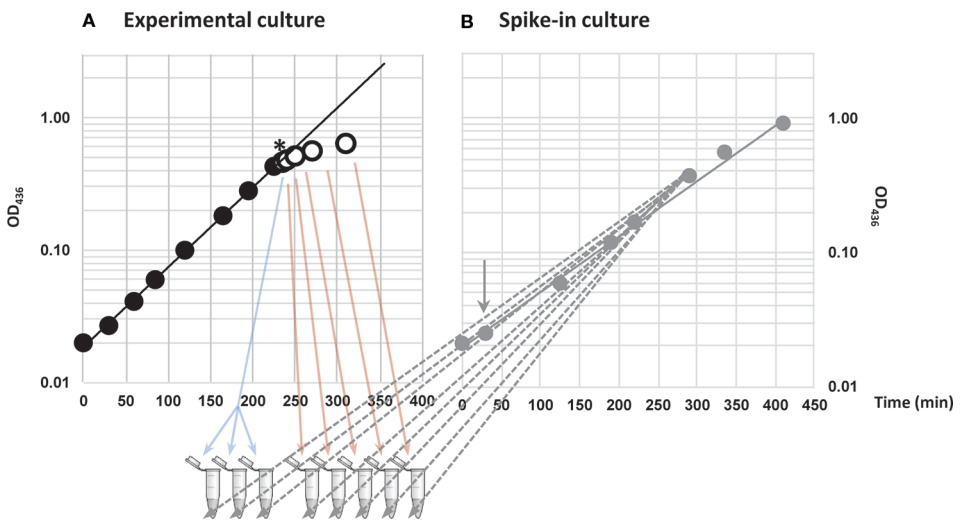

FIGURE 1 | Provoking amino acid starvation and addition of spike-in cells to samples. **(A)** Cells were grown in steady state (closed circles) before induction of isoleucine starvation by addition of 400 μg/ml L-valine (denoted by *). Three steady-state samples (blue arrows) were harvested as reference immediately before addition of L-valine and five samples in total were harvested during starvation up to 80 min after the addition of L-valine (open circles and red arrows). **(B)** In parallel, an *E. coli* strain that carried an inducible plasmid expressing a *Vibrio cholerae* small RNA (sRNA) (*qrr2*), antisense-*lacI* and an antibiotic marker (*bla*) (gray closed circles) was grown as a spike-in culture. Addition of inducer (1 mM IPTG) is indicated with a gray arrow. The samples collected in **(A)** were spiked-in with 1% optical density (OD) units of the spike-in culture of cells harboring the plasmid (gray dashed arrows).

purification was adjusted according to the samples' OD at the time of harvest, to ensure a constant ratio of spike-in cells to sampled cell mass (as measured by OD). We first assessed the spike-in method by calculating the ratio of spike-in reads to total reads. Thus, we could evaluate two parameters; i), how much the three steady-state replicate samples varied from each other and ii), whether the spike-in method indicated changes in total RNA levels during starvation. As seen in **Supplementary Figure S1A**, the ratio of plasmid reads to total reads of the three replicates taken during steady-state growth varied only by ~1%, indicating a high reproducibility of the data. In contrast, as starvation progressed within the 80-min time series, the ratio of plasmid reads/total reads increased, indicative of a decline in total RNA levels from the experimental samples, which is consistent with the net negative effect of ppGpp on the activity of RNA polymerase (Fiil et al., 1972; Sarubbi et al., 1988) and breakdown of rRNA (Zundel et al., 2009; Fessler et al., 2020) and tRNA (Svenningsen et al., 2017) upon starvation. However, the correlation deviated from the expected ratio at the 20-min time point with approximately 30% from the trend. The deviation is most likely due to erroneous sampling, which results in a surge in the ratio of ribosomal reads to spike-in reads at the 20-min time point (**Supplementary Figure S1B**). The surge in rRNA mid-starvation is highly unlikely to have a biological explanation, given the negative effect of amino acid starvation and ppGpp production on rRNA synthesis (Sands and Roberts, 1952; Cashel and Kalbacher, 1970). Therefore, we regarded the 20-min time point as an outlier and did not include it in the further analysis of the transcriptome. We then proceeded with normalizing the sequencing reads to the spike-in RNA (here designated RPKSP, Reads Per Kilobase of gene per 10 kilobase of spike-in, see *Materials and Methods*). For comparison we also normalized the sequencing reads (excluding reads mapping to the spike-in plasmid) using the standard method that only takes into account the sequencing depth and gene length (RPKM, Reads per Kilobase Million).

Hallmark Stringent Response Gene Regulation Is Captured With RNAseq

When *E. coli* experiences amino acid starvation, transcription of the protein synthesis machinery is adjusted within minutes to meet the lower demand for protein synthesis (Maaløe, 1979; Ryals et al., 1982; Nomura et al., 1984). This hallmark of the stringent response was clearly reflected in our transcriptomic data, shown in **Figure 2** by the mRNAs encoding ribosomal proteins and elongation factor Tu. **Figure 2** also shows that two extensively characterized promoters known to be activated by ppGpp, namely the *iraP* and *uspA* promoters (Nyström and Neidhardt, 1992; Bougdour et al., 2006; Vollmer and Bark, 2018), are up-regulated in this analysis. Thus, the ppGpp-mediated stringent response is activated upon L-valine-induced isoleucine starvation in our experiment, and the general trends are detected using both methods of data normalization (RPKSP and RPKM).

Ribosomal RNA Turnover Upon Isoleucine Starvation

Assessing the ribosomal RNAs, however, our spike-in-normalized data show that not only was the synthesis of rRNA down-regulated, but the levels of existing rRNA per OD unit of cells were substantially reduced upon isoleucine starvation. Specifically, after 80 min of starvation the levels of 16S and 23S rRNAs had decreased to approximately 70% of the pre-starvation level (**Figure 3A**, RPKSP). This behavior was only visible when we normalized the sequencing reads to levels of reads from the spike-in plasmid, and not to total reads (**Figure 3B**, RPKM). In agreement with the RPKSP-normalized data, northern blots showed that 16S and 23S rRNAs decayed to

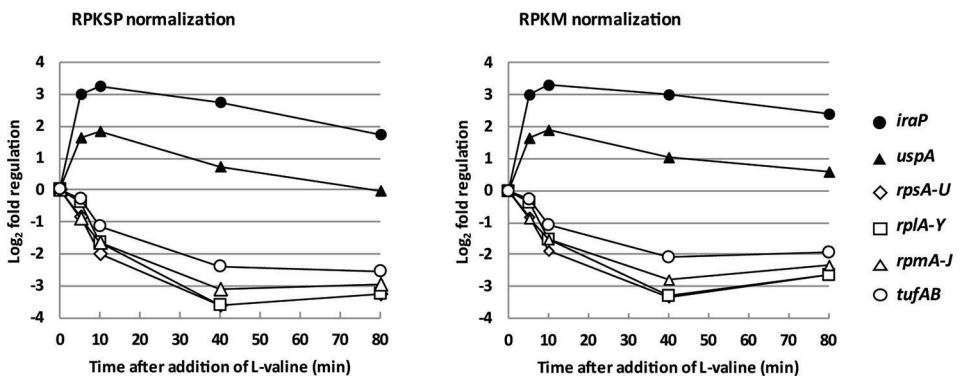

FIGURE 2 | The messenger RNA (mRNA) encoding protein components of the protein synthesis machinery are rapidly down-regulated, and known ppGpp-controlled stress response proteins are rapidly up-regulated upon isoleucine starvation. The average levels of ribosomal protein mRNA reads during isoleucine starvation plotted as $\log_2$-fold change relative to pre-starvation levels (*rpsA-U*: open diamonds, *rplA-Y*: open squares, and *rpmA-J*: open triangles) as well as the average mRNA levels of elongation factor EF-Tu (*tufAB*: circles) and mRNA levels of the anti-adaptor protein, *iraP* (closed circles), and universal stress protein A, *uspA* (closed triangles).

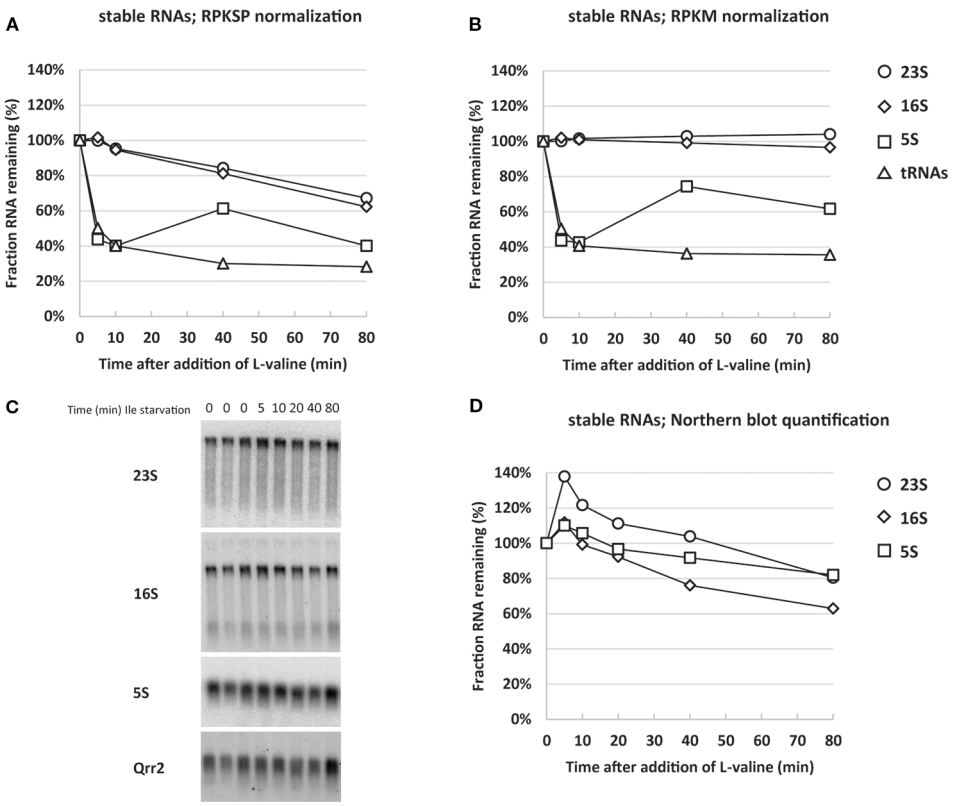

FIGURE 3 | Normalizing RNA sequencing reads to spike-in RNA reveals that stable RNA levels are substantially reduced during isoleucine starvation. The average reads of stable RNA (23S; circles, 16S; diamonds, 5S; squares and transfer RNAs (tRNAs); triangles) during isoleucine starvation are shown relative to the average pre-starvation levels normalized with two different methods: **(A)** by spike-in cells (RPKSP, Reads Per Kilobase of gene per 10 kilobase of spike-in) and **(B)** by total reads (RPKM, Reads per Kilobase Million). **(C)** A 1% agarose gel was used for electrophoresis of total RNA from three samples harvested in steady-state growth before induction of isoleucine (Ile) starvation (0 time points) and during starvation (5, 10, 20, 40, 80 min time points). The resulting blot was probed for 23S, 16S, 5S, and the spike-in-cell-specific RNA Qrr2 as indicated. **(D)** The levels of stable RNA (23S; circles, 16S; diamonds, 5S; squares) were quantified by normalizing to Qrr2 from the spike-in cells and shown relative to the average of the three RNA samples harvested prior to starvation. The quantified and normalized data originates from the blot in panel **(C)**.

approximately 60–80% of the pre-starvation level in the first 80 min after starvation (**Figures 3C, D**).

While there is good agreement between the two methods for 16S and 23S rRNA, there is a discrepancy in the quantification of 5S rRNA levels. In the northern blot analysis, 5S levels declined to approximately 80% after 80 min, whereas RNAseq reads indicate a decline to approximately 40% of the pre-starvation level. We suspect that the lower levels of 5S reads is likely a consequence of a higher number of mapping errors for short RNAs in the RNAseq pipeline, as we noticed a general increase in the variation between the triplicate steady-state samples for reads mapping to short genes (<0.2 kb) (**Supplementary Table S2**). As a further quantification control, we assessed the RNA samples by qRT-PCR for the levels of 5S. The qRT-PCR data verified the magnitude of 5S decline shown in the northern blots (**Supplementary Figure S2**), confirming that 5S was unreliably quantified in the RNAseq pipeline. Collectively, the RPKSP-normalized transcriptome, the northern blots, and the qRT-PCR assay, validate that rRNA levels decrease substantially during the early response to isoleucine starvation.

Transfer RNA Turnover Upon Isoleucine Starvation

A rapid reduction in tRNA levels upon L-valine-induced isoleucine starvation as well starvation for other amino acids was reported previously, but the kinetics of tRNA disappearance shown in **Figure 3** are much faster than expected from northern blot experiments (Svenningsen et al., 2017), regardless of the method of normalization. In addition, the concentration of tRNA is highly underestimated by the RNAseq method as a molar ratio of about 10 tRNAs per ribosome is expected (Dong et al., 1996), but we only detected 0.003 tRNA per rRNA by RNAseq during steady-state growth (**Supplementary Table S3**). The low detection of tRNAs is reportedly due to the difficulties in reverse transcription of the highly modified tRNA to cDNA (Motorin et al., 2007). While tRNA is quantified independently

of its modification status in northern blots, here it is not, and we therefore expect newly transcribed hypomodified tRNA to be overrepresented in the RNAseq analysis. This could explain why tRNA "disappears" fast (within 5 min of the onset of starvation; **Figure 3**) as transcription of tRNA genes is curtailed by the stringent response, so the pool of hypomodified tRNA is expected to decrease very fast upon starvation and enter the pool of poorly detected mature tRNA. Indeed, treatment of the starved culture with the transcription initiation inhibitor rifampicin, which terminates initiation of RNA synthesis, resulted in an additional decrease in tRNA-mapped reads down to just 2% of the pre-starvation level, supporting that very little mature tRNA was detected by RNAseq (**Figure 4**), while previous northern blot experiments showed at least 20% retention of tRNA 80 min after rifampicin treatment (Svenningsen et al., 2017). By contrast, the profile of rRNA levels per OD unit of culture remained nearly undisturbed during the rifampicin treatment (**Figure 4**).

Changes in the Size of the Total Messenger RNA Pool Upon Isoleucine Starvation

The whole-cell spike-in method in combination with RNAseq allowed us to estimate the kinetics of the reduction in the total mRNA pool during starvation (**Figure 5**). This estimate is unique in that it yields direct information on mRNA abundance per OD unit of bacterial culture under starvation relative to steady-state levels, whereas previous estimates were based on the change in synthesis rates of stable RNA relative to total RNA (R_S/R_T) during starvation (Ryals et al., 1982), or the addition of synthetic RNA spike-in after the preparation of sample RNA (see e.g., Schena et al., 1995; Gorochowski et al., 2019). It is well known that the promoter selectivity and the initiation frequency of RNA polymerase changes as a function of the ppGpp concentration (Kajitani and Ishihama, 1984; Sanchez-Vazquez et al., 2019) and that ppGpp switches RNA polymerase onto stress-related genes rather than genes for components of the translational apparatus (as illustrated in **Figure 2**). It has also been shown that the processivity of the RNA polymerase is negatively affected by the concentration of ppGpp (Kingston and Chamberlin, 1981; Kingston et al., 1981; Sørensen et al., 1994; Vogel and Jensen, 1994; Roghanian et al., 2015). In the present set of data (**Figure 5**) we can see how these effects of reduced RNA polymerase initiation frequency, processivity, and altered promoter selectivity combined to reduce the total mRNA pool to about 70% already after 10 min of starvation, and reduced it by half after 80 min (**Figure 5 and Supplementary Table S4**).

In summary, the spike-in methodology allowed us to quantify the change in the pools of rRNA and mRNA upon isoleucine starvation and subsequent rifampicin treatment, while tRNA could not be reliably quantified using this method. Northern blot analysis confirmed the decrease in rRNA shown by RPKSP normalization (**Figure 3C**). The underlying reason that RPKSP reveals this decrease while RPKM normalization does not, is that since the rRNA comprises ~85% of total RNA in the cell (on average 89% of the total reads in our samples), a decrease in rRNA will result in an almost equivalent decrease in the total RNA. Therefore, a normalization method that relies on sequencing depth will i) mask changes in very abundant rRNAs, ii) underestimate the magnitude of the change in

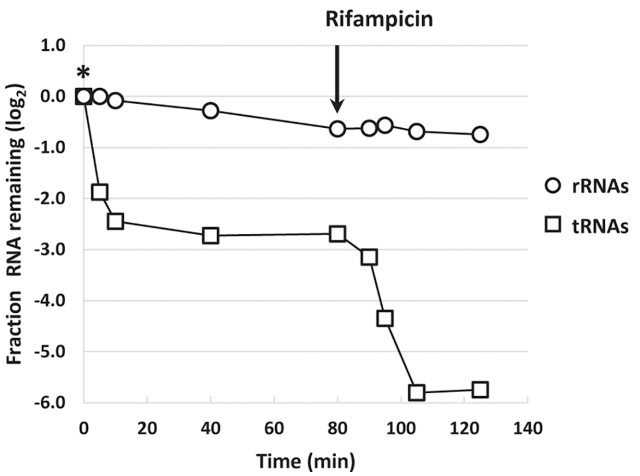

FIGURE 4 | Inhibition of transcription leads to diminished levels of transfer RNA (tRNA)-mapped reads. The sum of reads mapped to ribosomal RNA (rRNA) genes (open circles) and the sum of reads mapped to tRNA genes (open squares) during isoleucine starvation are shown as the log_2-fold change relative to their pre-starvation levels (L-valine addition denoted by *). Immediately after the last isoleucine starvation sample was harvested at 80 min, 300 µg/ml rifampicin was added to the culture to block transcription initiation. Samples were harvested 10, 15, 25, and 45 min after rifampicin addition.

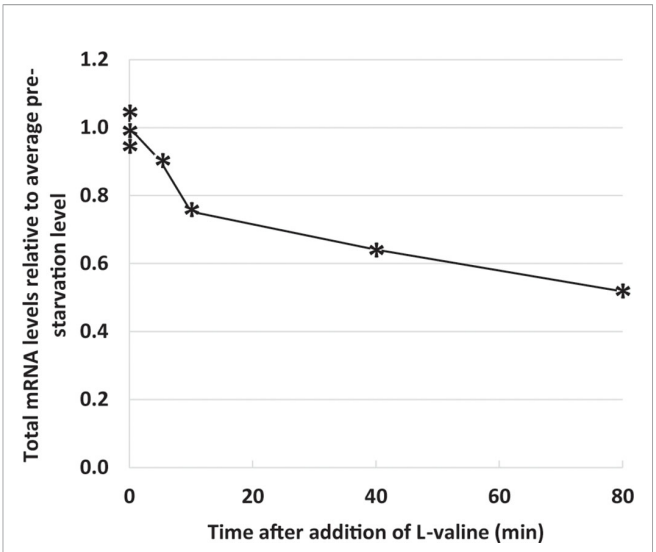

FIGURE 5 | Total messenger RNA (mRNA) levels decrease during starvation. Reads mapping to 129 noncoding RNAs [ribosomal RNA (rRNA), transfer RNA (tRNA), small regulatory RNAs, and the RNA component of RNase P] were removed from the filtered **Supplementary Data Sheet 1** to yield the mRNA dataset (**Supplementary Table S4**). The sum of reads mapped to the 3.919 mRNA genes during isoleucine starvation are plotted relative to the average of their pre-starvation levels.

RNAs that change in the same direction as the very abundant RNA, and iii) overestimate the magnitude of the change in RNAs that change in the opposite direction of the very abundant RNA.

Transcriptome-Wide Response to Amino Acid Starvation Induced by L-Valine

The sequencing results for individual genes are available in **Supplementary Data Sheet 1** (RPKSP-normalized) and **Supplementary Data Sheet 2** (RPKM-normalized) as the RNA abundance levels (normalized reads per gene) for each time point as well as the $\log_2$–fold difference in RNA abundance levels at each time point of starvation and rifampicin treatment, relative to the average of the three steady-state samples. **Supplementary Tables S5** and **S6** are alphabetic lists of RPKSP- and RPKM-normalized genes, that are up- or down-regulated more than two-fold at the 80-min starvation time point relative to the average of the three steady-state time points. Finally, **Supplementary Tables S8** and **S9** report the 100 RPKSP-normalized genes most strongly activated and repressed, respectively, upon isoleucine starvation.

The difference in the outcomes of the two normalization methods for L-valine-induced isoleucine starvation was assessed by plotting the 10 and 80 min time points relative to the average of the three steady-state samples (**Figures 6A, B**). We apply a two-fold regulatory threshold (Wren and Conway, 2006) to ease comparison between our data sets and the most relevant literature (Traxler et al., 2011; Sanchez-Vazquez et al., 2019). As shown in **Figure 6**, RPKM normalization underestimates the number of down-regulated genes in the stringent response compared to normalization to the spike-in reads (RPKSP). This effect is more pronounced as starvation progresses. RPKM normalization fails to detect 40% of the ≥ 2.0 fold down-regulated genes at 80 min post starvation, which are detected with RPKSP normalization (**Figure 6**). By contrast, RPKM overestimates the number of genes induced ≥2 fold by >40% at the 80 min time point, compared to RPKSP normalization (**Figure 6B**). While the number of genes that qualify for the ≥2 fold up- or down-regulation cut-off clearly differ substantially between the two normalization methods, we emphasize that the identity of the most strongly regulated genes is independent of the normalization method. Thus, the 970 genes that could be identified as ≥2-fold down-regulated in the RPKM-normalized data set despite the tendency for this method to overestimate gene expression late in starvation relative to the steady state, form the most strongly down-regulated subset of the 1642 genes that were identified as ≥2-fold down-regulated after application of the RPKSP correction, and *vice versa* for the up-regulated genes (**Supplementary Tables S5** and **S6**). From this result, it is evident that the method of normalization is critical for the interpretation of changes in RNA levels when cells experience a shift in growth condition. In the remainder of the manuscript, we therefore focus on the RPKSP-normalized data set as we analyze the transcriptomic response to amino acid starvation induced by L-valine.

Two Temporal Profiles Account for the Majority of Gene Expression Changes

To explore trends in the transcriptome response to isoleucine starvation in general terms, a principal component analysis (PCA) was carried out. PCA is a statistical procedure that uses linear transformations of the original data (relative abundance of each RNA at the seven time points, see *Materials and Methods*) to define a set of new, orthogonal variables that reduce the number of variables needed to describe the data set. We found that the first two components of the analysis account for 69% of the variability of the data. The temporal profile of these two components (PC1 and PC2) are shown in **Figures 7A, B** respectively, and PC1 and PC2 scores for each individual gene is provided in **Supplementary Table S7**. The PC1 vector, which accounts for 48% of the variability, describes genes that do not show variation among the three steady-state samples, change abruptly in response to the addition of L-valine, and remain at the new level throughout the duration of the starvation. An example of a gene with a high positive PC1 score is *uspB*, encoding the universal stress protein B, which is known to be induced by starvation (Farewell et al., 1998). Like PC1, the PC2 vector describes genes that do not vary among the three steady-state samples, and change abruptly in response to the addition of L-valine. But in contrast to PC1, the RNA level for genes with a high positive PC2 score show a surge at the 5 min time-point followed by a drop as starvation continues. An example of such a gene is *crp*, encoding the cAMP-binding global transcriptional regulator CRP. **Figure 7C** shows all 4,048 genes in the

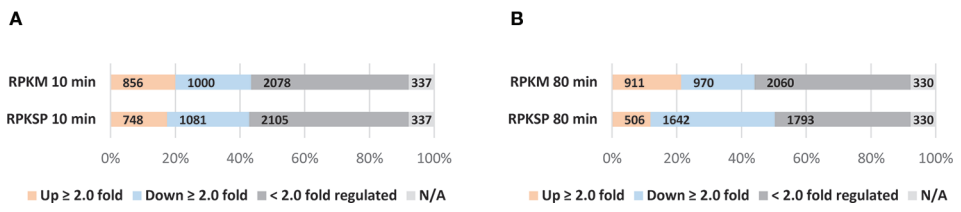

FIGURE 6 | Applying the spike-in methodology alters the proportions of up- and down-regulated genes observed during isoleucine starvation. Normalization was done by two different methods; RPKM (Reads Per Kilobase Million) and RPKSP (Reads Per Kilobase of gene per 10 kilobase of spike-in). RNAs that showed >2-fold change at **(A)** 10 min and **(B)** 80 min after starvation relative to the average of the three samples harvested immediately before isoleucine starvation are denoted in orange bars (≥ 2.0 fold up-regulated) and blue bars (≥ 2.0 fold down regulated). Genes that were < 2.0-fold regulated genes are shown in dark gray bars. N/A: not included in the analysis either because the reads in the triplicate steady state samples were filtered out (see *Materials and Methods*), or gene reads were not detected in either steady-state growth or starvation (light gray bars).

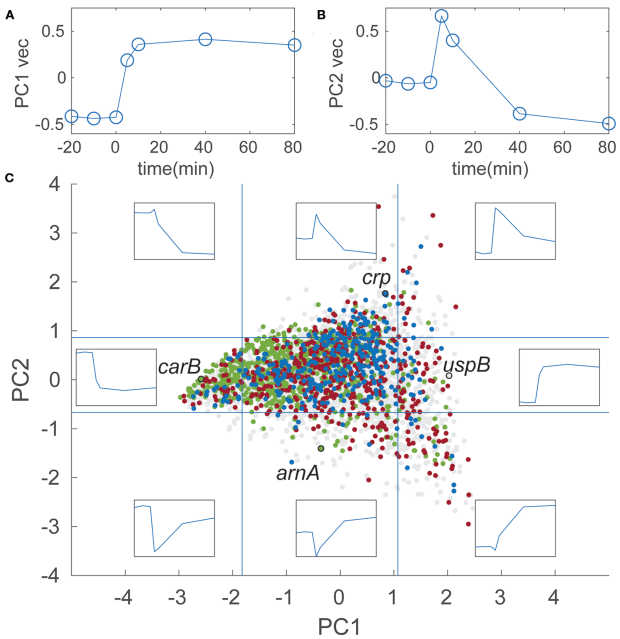

FIGURE 7 | Two distinct temporal profiles account for the majority of gene expression changes in the stringent response transcriptome. **(A)** Temporal profile of the PC1 vector. The three steady-state samples are artificially displayed between time –20 and 0 to better illustrate the shape of the vector. Units on the y-axis are arbitrary units of normalized RNA levels, as described in *Materials and Methods*. **(B)** Temporal profile of the PC2 vector; **(C)** 4,048 genes plotted according to their PC1 and PC2 values. Inserts show temporal profiles at the corresponding coordinates, e.g., the top left insert shows the profile for PC1 = –3 and PC2 = 3. Vertical blue lines indicate the cut-off values for genes in the subset with 10% highest and lowest PC1 values used for enrichment analysis. Horizontal blue lines indicate the same for the PC2 values. Each data point corresponds to a gene and is colored according to the parent GO term it belongs to. Since some genes belonged to more than one category, the coloring was layered so that the final graph displays the smallest category the data point belongs to: all data points (4,048 points) light gray; biosynthetic process (975 points) green; response to stress (578 points) red; regulation of transcription, DNA-templated (349 points), blue. The points corresponding to the genes mentioned in the main text are highlighted with black open circles and labeled with the gene names.

transcriptome analysis plotted according to their PC1 and PC2 scores (colored dots). The temporal profiles for selected values of PC1 and PC2 are shown as eight inserts at the corresponding positions on the graph. For example, the insert at position (−4;0) depicts a gene with a temporal profile that is strongly negatively correlated with the PC1 profile shown in **Figure 7A**. An example of such a gene is *carB*, encoding a component of carbamoyl phosphate synthetase, which is involved in arginine biosynthesis from ornithine (see also **Figure 8C**). The gene *arnA* encoding a key enzyme in outer membrane lipid A modification (Williams et al., 2005), is an example of a gene with a negative correlation to the PC2 profile. Specifically, *arnA* mRNA was quite abundant during steady-state growth, dropped 15-fold at 5 minutes after starvation, and returned to steady-state levels after 40 min of starvation. The temporal profile of *uspB*, *crp*, *carB*, and *arnA* were confirmed by qRT-PCR and show similar relative expression profiles as the RPKSP-normalized sequencing data (**Supplementary Figure S3**).

To explore whether general trends could be discerned in terms of the biological processes associated with temporal profiles defined in the PCA, we used the enrichment tool available on the EcoCyc webserver (Keseler et al., 2017) to identify gene ontology (GO) terms that were statistically over-represented in the four subsets of genes that scored among the 10% highest or lowest for PC1 or PC2 (see *Materials and Methods*). Among the tens to hundreds of GO terms that were significantly enriched in each subset, we focused on broad categories (parent GO terms) to highlight the general trends in the dataset, rather than focus on specific metabolic pathways or regulons. The 10% of genes that had the highest PC1 scores were enriched (p-value $3*10^{-7}$) for the parent GO term "response to stress" (**Figure 7C**, red dots). By contrast, the 10% of genes that had the lowest PC1 scores were highly enriched (p-value $7*10^{-13}$) for the broad GO term "biosynthetic process" (**Figure 7C**, green dots). Meanwhile, the 10% of genes that had the highest PC2 scores were enriched (p-value $2*10^{-4}$) for "regulation of transcription, DNA-templated" (**Figure 7C**, blue dots), while those with the 10% lowest PC2 scores did not yield a significantly enriched broad category. While there are naturally many outliers within these broad categories, this analysis illustrates that general temporal profiles can be recognized in the stringent response that distinguish biosynthesis genes which generally remain down-regulated during starvation (low PC1), and stress response genes which generally remain up-regulated (high PC1), from the transcriptional regulators whose expression spikes during the growth transition followed by a recovery period (high PC2).

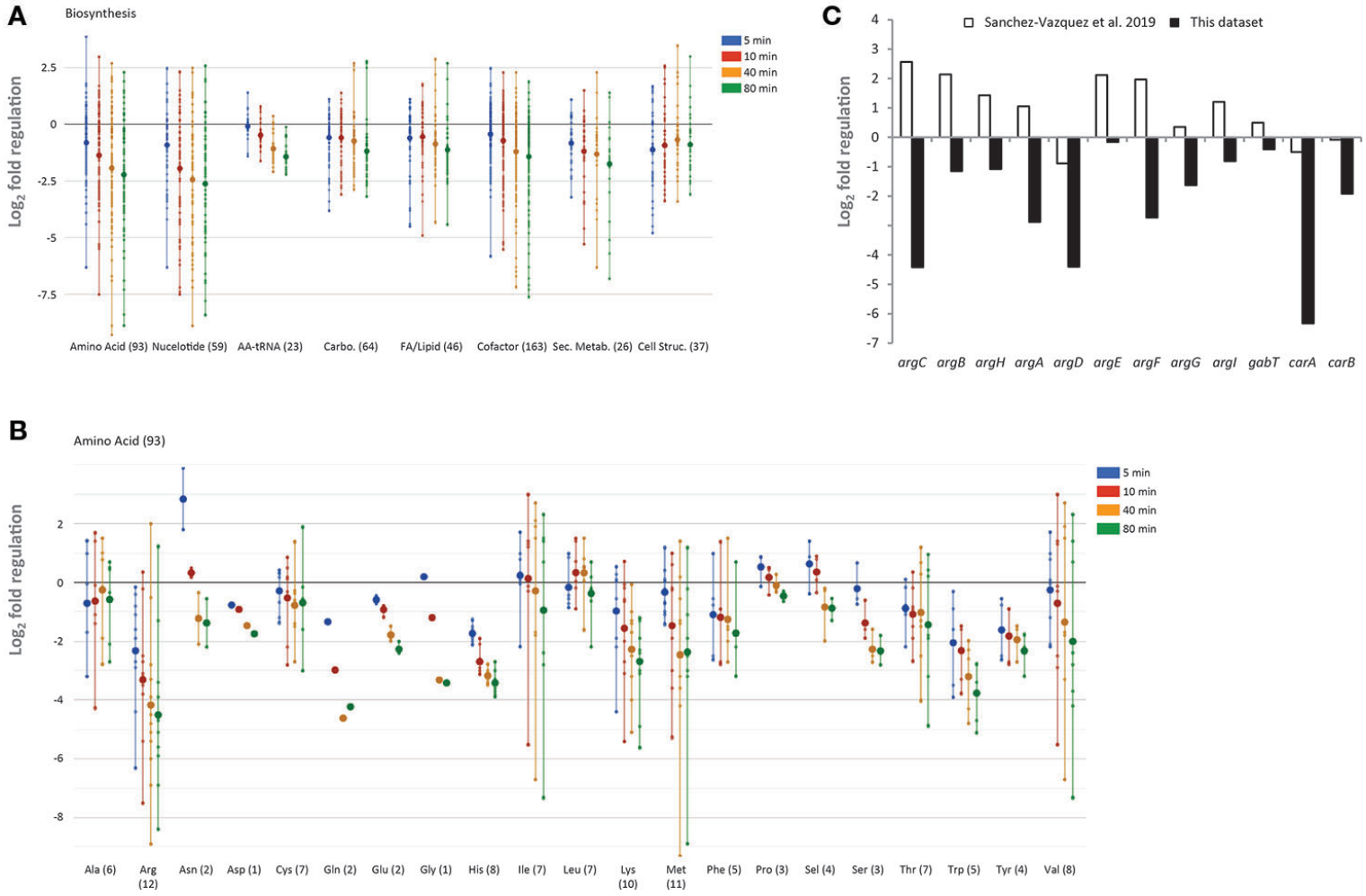

FIGURE 8 | Regulation of biosynthetic genes upon L-valine-induced isoleucine starvation. **(A)** Temporal profile of biosynthetic genes as displayed by the "Omics Dashboard" tool (Paley et al., 2017). Each gene in a class is represented by a dot for each of the four time-points during starvation (5, 10, 40, and 80 min) The position on the y-axis indicates the log2-fold change in RNA level relative to the average of the three steady-state samples. The large circles represent the average change of genes in each category at each time point. Genes are classified as belonging to amino acid biosynthesis, nucleoside and nucleotide biosynthesis, carbohydrate biosynthesis, fatty acids and other lipids biosynthesis, aminoacyl-tRNA synthetases, biosynthesis of small molecules (co-factors, prosthetic groups, electron carriers, and vitamins) that participate in enzyme reactions, biosynthesis of secondary metabolites, and biosynthesis of cell structural elements (mainly cell wall). A gene can be assigned to more than one class. The number of genes in each class is indicated in parentheses. **(B)** Temporal profile of biosynthetic genes belonging to the sub-system amino acid biosynthesis. **(C)** Change in messenger RNA (mRNA) level of genes in the arginine biosynthesis pathway 5 min after addition of L-valine, relative to the average of the three steady-state samples (black bars). White bars show the relative change of the same mRNAs 5 min after ectopic induction of high levels of ppGpp produced by a constitutive variant of *relA* as reported by Sanchez-Vazquez et al. (2019).

Downregulation of Biosynthesis

The "Omics Dashboard" software tool available on the EcoCyc webserver (Paley et al., 2017) was used to further explore the transcriptomic changes that occurred in response to L-valine addition. The tool combines data on the expression level of individual genes into a hierarchy of cellular systems and subsystems. As expected in response to amino acid starvation, and as indicated in the PCA analysis, genes responsible for the major biosynthetic processes (e.g., nucleotide, carbohydrate, fatty acid, lipid, and aminoacyl-tRNA synthesis) are generally down-regulated (**Figure 8A**). Amino acid biosynthesis genes are reportedly up-regulated under the stringent response, which could help *E. coli* overcome starvation (Cashel et al., 1996), and this has been observed to varying extents in previous transcriptome-wide analyses (Cashel et al., 1996; Durfee et al., 2008; Traxler et al., 2008; Sanchez-Vazquez et al., 2019). We found that amino acid biosynthetic genes were generally down-regulated in response to L-valine (**Figure 8B**). For example, while seven of the 12 genes ascribed to arginine biosynthesis were up-regulated within 5 min of ppGpp production in a recent study where a constitutively active RelA variant was induced to produce ppGpp in the absence of starvation (Sanchez-Vazquez et al., 2019), none of the arginine biosynthetic genes were up-regulated in our study (**Figure 8C**). We expect that the key difference between the two experiments is that all amino acids were supplied in the growth medium in the study conducted by Sanchez-Vazquez and coworkers, so at the outset of the experiment, arginine biosynthetic operons would be repressed by the arginine-bound ArgR repressor (Caldara et al., 2006). What is then measured is a positive regulatory effect of ppGpp on some of these repressed promoters. In our study, the growth medium did not contain any amino acids prior to the addition of L-valine, so the amino acid biosynthetic operons were already de-repressed when ppGpp production was induced by L-valine-mediated isoleucine starvation. We suspect that the reduced rate of protein synthesis that occurs upon isoleucine starvation results in a build-up of the residual amino acids, including arginine, which would lead to repression of the arginine biosynthesis pathway by arginine-bound ArgR.

Differential Response of the RpoS and Lrp Regulons

In most bacterial systems, the stringent response includes a robust general stress response mediated by the stationary phase sigma factor RpoS (σ^S or σ^{38}), which is regulated at the level of transcription, translation as well as at the level of protein stability both directly and indirectly by ppGpp (Lange and Hengge-Aronis, 1994; Landini et al., 2014). The RpoS regulon has been studied extensively and is known to control >140 genes in response to diverse stress conditions (Lacour and Landini, 2004; Weber et al., 2005), including isoleucine starvation (Traxler et al., 2011). In the study by Traxler and coworkers, they analogously applied isoleucine starvation on the conditional auxotrophic *E. coli* K-12 strain, but contrasting our experiments, the cells in their experimental set-up gradually exhausted isoleucine in media containing all other amino acids. They show that the levels of ppGpp calibrate and co-regulate the RpoS-dependent stress response and the Lrp-dependent regulon (leucine responsive protein), which mostly includes genes for metabolic enzymes. The Lrp-dependent response occurred prior to and at lower ppGpp concentrations than the RpoS-dependent response (Traxler et al., 2011). We employed the definitions of the RpoS and Lrp regulons used by Traxler et al., and investigated the isoleucine starvation response of these two regulons in our experimental set-up. As seen in **Figure 9A**, the majority of the RpoS-dependent genes are induced after 40 min of L-valine-mediated isoleucine starvation, in agreement with the slow but robust induction of the regulon reported in the previous study. However, under the condition tested here, *E. coli* did not significantly induce the Lrp regulon apart from genes involved in alanine metabolism (*dadAX*) (**Figure 9B**). In fact, the *lrp* mRNA itself was three-fold down-regulated at the end-point of the starvation. In line with this finding, the small regulatory RNA GcvB was among the top 10 up-regulated transcripts in our experiment (**Supplementary Table S8**), and GcvB is known to regulate the *lrp* mRNA negatively (Holmqvist et al., 2012; Lee and Gottesman, 2016; Lalaouna et al., 2019). It is unknown to us how valine-induced isoleucine starvation could trigger high expression levels of GcvB but we suggest that the induction of GcvB could be the main reason for the missing Lrp response in this particular experimental set-up.

A Specific Transcriptional Response to L-Valine-Induced Isoleucine Starvation

The only gene which was up-regulated more strongly than GcvB 5 min after starvation was *alaE*, encoding an L-alanine exporter (Hori et al., 2011), which showed an average increase of transcript levels during starvation >300-fold compared to pre-starvation levels (**Supplementary Table S8**). *alaE* transcription was recently shown to be positively regulated by ppGpp (Sanchez-Vazquez et al., 2019), supporting the up-regulation seen here. Although not included in the Lrp regulon defined in (Traxler et al., 2011), *alaE* is also predicted to be up-regulated by Lrp. Similarly, *dadA* and *dadX* which are identified here as the only clearly up-regulated genes in the curtailed Lrp-regulon (**Figure 9B**) were identified to be transcriptionally activated by ppGpp in the same study (Sanchez-Vazquez et al., 2019). We suggest that the prominent up-regulation of *alaE, dadA, dadX,* and *gcvB* results directly from cellular metabolic consequences of the addition of L-valine rather than the resulting starvation for isoleucine. The physiological role of AlaE is to export L-alanine (and possibly alanine dipeptide) to avoid intracellular toxic-level accumulation of L-alanine (Kim et al., 2015). The excess L-valine, together with pyruvate as substrate, can be converted to L-alanine by *avtA,* one of three major alanine-synthesizing transaminases in *E. coli* (Hori et al., 2011). The *avtA* mRNA was not up-regulated during starvation, but it was highly expressed during unrestricted growth under our conditions (**Supplementary Data Sheet 1**), suggesting that the transaminase protein it encodes is abundant at the onset of starvation. According to this model, overabundant levels of L-alanine is exported out from the cell by AlaE. In addition, surplus

L-alanine can be converted by the alanine racemase *dadX* to D-alanine (Wild et al., 1985), which in turn is the substrate for the D-amino acid dehydrogenase *dadA* in the inner membrane to yield ammonium and pyruvate (Franklin and Venables, 1976). This further fuels the conversion of L-valine to L-alanine (**Figure 10**). Evidently, excess L-valine gave rise to high levels of L-alanine that is countermeasured by upregulating the mRNA encoding the alanine exporter, clearly envisaged in this transcriptome. Moreover, elevated levels of D-alanine in the cells might be utilized and directed to cell wall synthesis. Some cell structure biosynthetic genes were de-repressed as starvation progressed (**Figure 8A**), especially genes involved in UDP-MurNAc-pentapeptide biosynthesis (e.g., *ddlB* and *murD/F*) and peptidoglycan maturation (*mtgA*) (**Supplementary Table S10**).

DISCUSSION

The stringent response to amino acid starvation is in many respects a model system for studies of bacterial stress responses, and has been the subject of intense study for decades, including several transcriptome-wide studies (Durfee et al., 2008; Traxler et al., 2008; Traxler et al., 2011). Here, we combined RNAseq with spike-in-cell normalization of the sequencing depth to obtain an adjusted view of the stringent response that is independent of any assumptions about the total RNA content of the cells. The methodology allowed us to quantify the changes in total rRNA and total mRNA per OD unit of bacterial culture over the first 80 min of starvation for isoleucine. In accordance with other reports (Ben-Hamida and Schlessinger, 1966; Jacobson and Gillespie, 1968; Maruyama and Mizuno, 1970; Zundel et al., 2009; Piir et al., 2011; Fessler et al., 2020), we find that the stability of rRNA is compromised upon nutrient starvation, resulting in a drop in rRNA per OD unit to 70% of the pre-starvation level within the first 80 min of an amino acid starvation. Because rRNA constitutes the vast majority of cellular RNA, this drop affects the quantification of all other RNA species in the cell if the RNAseq data is normalized solely to the sequencing depth of the samples in the conventional way (referred to here as RPKM-normalization, **Figure 6**). One important outcome of our work is therefore that ~40% more mRNAs are down-regulated, and ~40% fewer are up-regulated by more than two-fold, compared to what a conventional RNAseq study would suggest. We remark that the problem associated with normalizing solely to sequencing depth is not solved by depletion of rRNA prior to RNA-seq, because the rRNA-depleted transcriptome also is subject to the

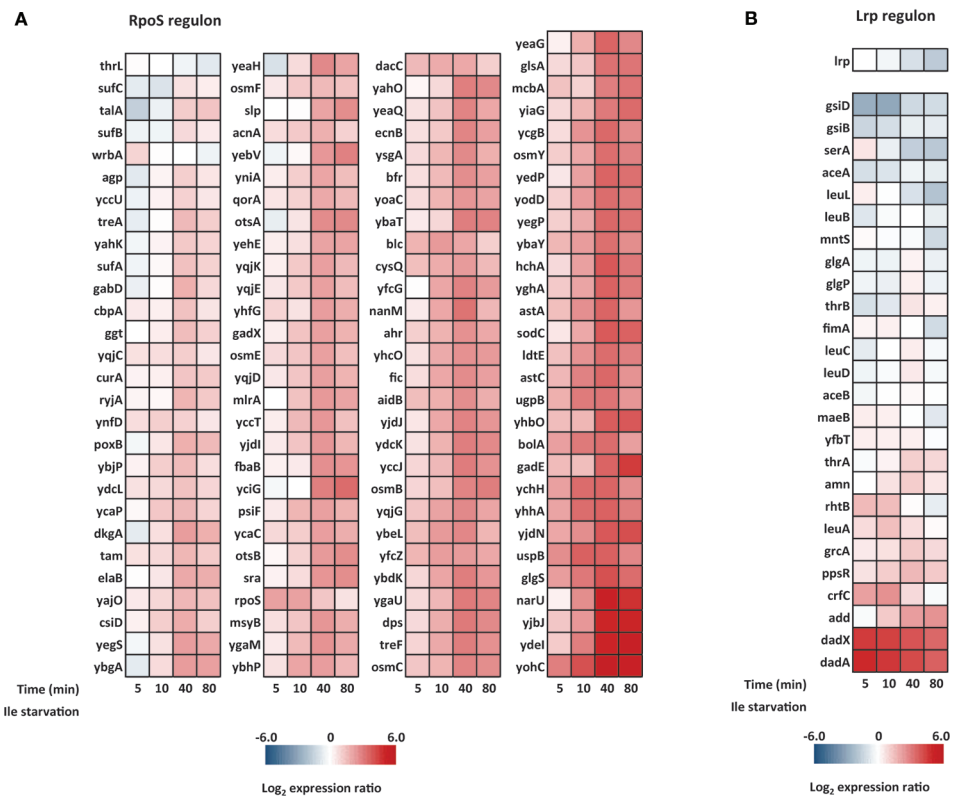

FIGURE 9 | Heat maps of the RpoS and Lrp regulons in cells starved for isoleucine. The time points into isoleucine starvation is indicated below each heat map. **(A)** The ppGpp/RpoS regulon and **(B)** the ppGpp/Lrp regulon. Genes listed in each regulon were defined in a previous study by Traxler and coworkers (Traxler et al., 2011).

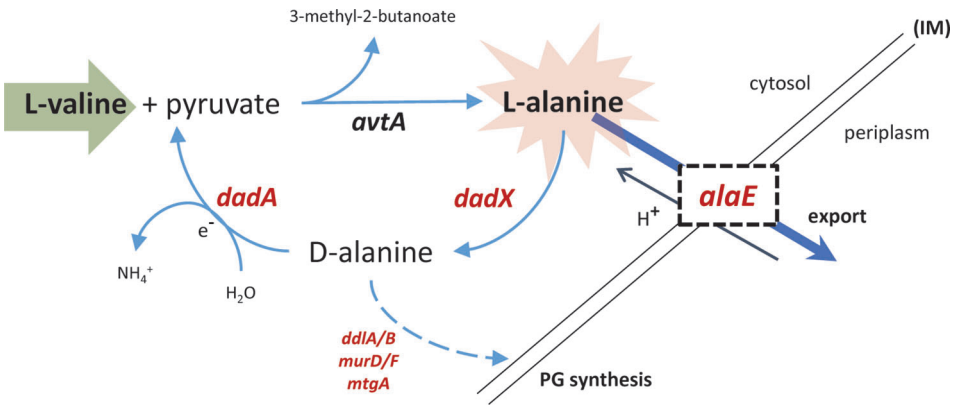

FIGURE 10 | Export pathway for excess L-valine. The synthesis of L-alanine requires pyruvate and is catalyzed by the valine-pyruvate aminotransferase *avtA*. The predominant alanine racemase in the cell, *dadX*, is degradative and catalyzes the conversion of L-alanine to D-alanine. The respiratory-chain-associated *dadA* can further catabolite D-alanine to pyruvate which can enter the central metabolism, or in this case be used as substrate to further convert high concentrations of L-valine to L-alanine. *alaE*, an alanine-proton antiporter, facilitates the export of L-alanine out from the cytosol in exchange of a proton. *ddlA/B, murD/F,* and *mtgA* are all involved in biosynthesis of cell wall structural elements. PG, peptidoglycan; IM, inner membrane). Up-regulated genes in the pathway are represented in red.

transcriptional consequences of a change in growth conditions. For example, the activity of RNA polymerase is reduced at elevated ppGpp levels, giving rise to lower RNA levels and the RNA chain growth rate is decreased (Kingston and Chamberlin, 1981; Kingston et al., 1981; Sørensen et al., 1994; Vogel and Jensen, 1994; Roghanian et al., 2015).

This study highlights that although the stringent response of *E. coli* to amino acid starvation has a set of defining characteristics, most notably a surge in ppGpp levels and reduced transcription of genes encoding the protein synthesis machinery, the particular growth conditions employed give notable differences in the transcriptome-wide response at the detailed resolution of an RNAseq experiment. Most notably, in contrast to a previous study (Traxler et al., 2011) the extensive Lrp regulon was not activated in response to isoleucine limitation in this study, and amino acid biosynthesis was not generally induced although many operons encoding amino acid biosynthesis genes are activated in response to ppGpp under other growth conditions (Sanchez-Vazquez et al., 2019). We used principal component analysis combined with enrichment analyses to identify broad functional classes of genes that responded similarly to the growth transition. Besides these, the data set also contains many smaller gene categories that will be of interest to specific research sub fields. For example, mRNA of the conserved BluR-repressed operon *ycgZ-ymgABC*, which were completely repressed during steady-state growth, were among the genes most strongly up-regulated upon starvation (**Supplementary Table S8**), suggesting an unidentified regulatory mechanism that is unrelated to the known BluR signals; blue light and low temperature, for the YcgZ regulator of OmpF porin expression and the Ymg biofilm modulators (Tschowri et al., 2012; Duval et al., 2017). Finally, at least four (*alaE*, *gcvB*, *dadA*, *dadX*) of the 20 genes that respond most strongly in our dataset are likely responding to the sudden addition of L-valine rather than the starvation for isoleucine (**Figure 10**).

The methodology also allowed us to quantify changes in total mRNA levels under starvation (**Figure 5**), and the result underlines that overall mRNA production is substantially reduced upon amino acid starvation. The rapid reduction of the total mRNA pool demonstrated here supports the model previously proposed to explain why the initial surge in ppGpp levels upon amino acid starvation levels off on a timescale of a few minutes, namely that the initial surge of ppGpp in response to the onset of starvation should taper off due to a reduction in the number of RelA-associated stalled ribosomes, resulting from the reduced availability of mRNA (Sørensen et al., 1994; Tian et al., 2016).

DATA AVAILABILITY STATEMENT

The raw reads generated for this study have been deposited in NCBI's Gene Expression Omnibus with the accession ID: GSE136753.

AUTHOR CONTRIBUTIONS

BG, MS, and SLS conceived and designed the study. BG, AB, MF performed experiments. BG, SAS, NM, MS, and SLS analyzed the data. BG, MS, and SLS wrote the paper.

FUNDING

We acknowledge funding of this project by the Danish National Research Foundation (DNRF120) and the Independent Research Fund Denmark (8049-00071B and 8021-00280A).

ACKNOWLEDGMENTS

We thank Marit Warrer for excellent technical assistance, Mette Kongstad and Thomas Søndergaard Stenum for enlightening discussions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00144/full#supplementary-material

REFERENCES

Abdi, H., and Williams, L. J. (2010). Principal component analysis. *Wiley Interdiscip. Rev.: Comput. Stat.* 2 (4), 433–459. doi: 10.1002/wics.101

Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Čech, M., et al. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* 46 (W1), W537–W544. doi: 10.1093/nar/gky379

Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31 (2), 166–169. doi: 10.1093/bioinformatics/btu638

Artsimovitch, I., Patlan, V., Sekine, S., Vassylyeva, M. N., Hosaka, T., Ochi, K., et al. (2004). Structural basis for transcription regulation by alarmone ppGpp. *Cell* 117 (3), 299–310. doi: 10.1016/S0092-8674(04)00401-5

Baracchini, E., and Bremer, H. (1987). Determination of synthesis rate and lifetime of bacterial mRNAs. *Anal. Biochem.* 167 (2), 245–260. doi: 10.1016/0003-2697(87)90160-6

Bartholomäus, A., Fedyunin, I., Feist, P., Sin, C., Zhang, G., Valleriani, A., et al. (2016). Bacteria differently regulate mRNA abundance to specifically respond to various stresses. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 374 (2063), 20150069. doi: 10.1098/rsta.2015.0069

Ben-Hamida, F., and Schlessinger, D. (1966). Synthesis and breakdown of ribonucleic acid in Escherichia coli starving for nitrogen. *Biochim. Biophys. Acta* 119 (1), 183–191. doi: 10.1016/0005-2787(66)90049-9

Bernstein, J. A., Khodursky, A. B., Lin, P. H., Lin-Chao, S., and Cohen, S. N. (2002). Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent DNA microarrays. *Proc. Natl. Acad. Sci. U. S. A* 99 (15), 9697–9702. doi: 10.1073/pnas.112318199

Bougdour, A., Wickner, S., and Gottesman, S. (2006). Modulating RssB activity: IraP, a novel regulator of σS stability in Escherichia coli. *Genes Dev.* 20 (7), 884–897. doi: 10.1101/gad.1400306

Bremer, H., and Dennis, P. (1996). "Modulation of chemical composition and other parameters of the cell by growth rate.," in *Escherichia coli and Salmonella: Cellular and Molecular Biology, edn 2 ed.* Eds. F. C. C.I.R. Neidhardt, J. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. Reznikoff, M. Schaechter, H. E. Umbarger and M. Riley (Washington, DC: American Society for Microbiology), 1553–1569.

Caldara, M., Charlier, D., and Cunin, R. (2006). The arginine regulon of Escherichia coli: whole-system transcriptome analysis discovers new genes and provides an integrated view of arginine regulation. *Microbiology* 152 (11), 3343–3354. doi: 10.1099/mic.0.29088-0

Cashel, M., and Kalbacher, B. (1970). The control of ribonucleic acid synthesis in Escherichia coli. V. Characterization of a nucleotide associated with the stringent response. *J. Biol. Chem.* 245 (9), 2309–2318.

Cashel, M., Gentry, D. R., Hernandez, V. J., and Vinella, D. (1996). "The Stringent Response," in *Escherichia coli and Salmonella: cellular and molecular biology, 2. ed.* Ed. F. C. Neidhardt. (Washington: ASM Press), 1458–1496.

Croucher, N. J., and Thomson, N. R. (2010). Studying bacterial transcriptomes using RNA-seq. *Curr. Opin. Microbiol.* 13 (5), 619–624. doi: 10.1016/j.mib.2010.09.009

Deutscher, M. P. (2003). Degradation of stable RNA in bacteria. *J. Biol. Chem.* 278 (46), 45041–45044. doi: 10.1074/jbc.R300031200

Dong, H., Nilsson, L., and Kurland, C. G. (1996). Co-variation of tRNA abundance and codon usage in Escherichia coli at different growth rates. *J. Mol. Biol.* 260 (5), 649–663. doi: 10.1006/jmbi.1996.0428

Durfee, T., Hansen, A. M., Zhi, H., Blattner, F. R., and Jin, D. J. (2008). Transcription profiling of the stringent response in Escherichia coli. *J. Bacteriol.* 190 (3), 1084–1096. doi: 10.1128/JB.01092-07

Duval, V., Foster, K., Brewster, J., and Levy, S. B. (2017). A novel regulatory cascade involving BluR, YcgZ, and lon controls the expression of escherichia coli OmpF Porin. *Front. Microbiol.* 8, 1148 1609–1625. doi: 10.3389/fmicb.2017.01148

Edgar, R., Domrachev, M., and Lash, A. E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30 (1), 207–210. doi: 10.1093/nar/30.1.207

Farewell, A., Kvint, K., and Nystrom, T. (1998). uspB, a new sigmaS-regulated gene in Escherichia coli which is required for stationary-phase resistance to ethanol. *J. Bacteriol.* 180 (23), 6140–6147.

Fessler, M., Gummesson, B., Charbon, G., Svenningsen, S. L., and Sorensen, M. A. (2020). Short-term kinetics of rRNA degradation in Escherichia coli upon starvation for carbon, amino acid, or phosphate. *Mol. Microbiol.* 1–13. doi: 10.1111/mmi.14462

Fiil, N. P., von Meyenburg, K., and Friesen, J. D. (1972). Accumulation and turnover of guanosine tetraphosphate in Escherichia coli. *J. Mol. Biol.* 71 (3), 769–783. doi: 10.1016/S0022-2836(72)80037-8

Franklin, F. C., and Venables, W. A. (1976). Biochemical, genetic, and regulatory studies of alanine catabolism in Escherichia coli K12. *Mol. Gen. Genet.* 149 (2), 229–237. doi: 10.1007/bf00332894

Gorochowski, T. E., Chelysheva, I., Eriksen, M., Nair, P., Pedersen, S., and Ignatova, Z. (2019). Absolute quantification of translational regulation and burden using combined sequencing approaches. *Mol. Syst. Biol.* 15 (5), e8719. doi: 10.15252/msb.20188719

Gummesson, B., Lovmar, M., and Nystrom, T. (2013). A proximal promoter element required for positive transcriptional control by guanosine tetraphosphate and DksA protein during the stringent response. *J. Biol. Chem.* 288 (29), 21055–21064. doi: 10.1074/jbc.M113.479998

Haseltine, W. A., and Block, R. (1973). Synthesis of guanosine tetra- and pentaphosphate requires the presence of a codon-specific, uncharged transfer ribonucleic acid in the acceptor site of ribosomes. *Proc. Natl. Acad. Sci. U. States America* 70 (5), 1564–1568. doi: 10.1073/pnas.70.5.1564

Holmqvist, E., Unoson, C., Reimegard, J., and Wagner, E. G. (2012). A mixed double negative feedback loop between the sRNA MicF and the global regulator Lrp. *Mol. Microbiol.* 84 (3), 414–427. doi: 10.1111/j.1365-2958.2012.07994.x

Hori, H., Yoneyama, H., Tobe, R., Ando, T., Isogai, E., and Katsumata, R. (2011). Inducible L-alanine exporter encoded by the novel gene ygaW (alaE) in Escherichia coli. *Appl. Environ. Microbiol.* 77 (12), 4027–4034. doi: 10.1128/AEM.00003-11

Hornung, B. V. H., Zwittink, R. D., and Kuijper, E. J. (2019). Issues and current standards of controls in microbiome research. *FEMS Microbiol. Ecol.* 95 (5). doi: 10.1093/femsec/fiz045

Jacobson, A., and Gillespie, D. (1968). Metabolic events occurring during recovery from prolonged glucose starvation in Escherichia coli. *J. Bacteriol.* 95 (3), 1030–1039. doi: 10.1128/JB.95.3.1030-1039.1968

Jones, M. B., Highlander, S. K., Anderson, E. L., Li, W., Dayrit, M., Klitgord, N., et al. (2015). Library preparation methodology can influence genomic and functional predictions in human microbiome research. *Proc. Natl. Acad. Sci. U. S. A.* 112 (45), 14024–14029. doi: 10.1073/pnas.1519288112

Ju, X., Li, D., and Liu, S. (2019). Full-length RNA profiling reveals pervasive bidirectional transcription terminators in bacteria. *Nat. Microbiol.* 4 (11), 1907–1918. doi: 10.1038/s41564-019-0500-z

Kajitani, M., and Ishihama, A. (1984). Promoter selectivity of *Escherichia coli* RNA polymerase. Differential stringent control of the multiple promoters from ribosomal RNA and protein operons. *J. Biol. Chem.* 259, 1951–1957.

Karp, P., Weaver, D., Paley, S., Fulcher, C., Kubo, A., Kothari, A., et al. (2014). The ecocyc database. *EcoSal Plus*. doi: 10.1128/ecosalplus.ESP-0009-2013

Keseler, I. M., Mackie, A., Santos-Zavaleta, A., Billington, R., Bonavides-Martinez, C., Caspi, R., et al. (2017). The EcoCyc database: reflecting new knowledge about Escherichia coli K-12. *Nucleic Acids Res.* 45 (D1), D543–D550. doi: 10.1093/nar/gkw1003

Kim, S., Ihara, K., Katsube, S., Hori, H., Ando, T., Isogai, E., et al. (2015). Characterization of the l-alanine exporter AlaE of Escherichia coli and its potential role in protecting cells from a toxic-level accumulation of l-alanine and its derivatives. *Microbiologyopen* 4 (4), 632–643. doi: 10.1002/mbo3.269

Kingston, R. E., and Chamberlin, M. J. (1981). Pausing and attenuation of *in vitro* transcription in the rrnB operon of E. coli. *Cell* 27 (3 Pt 2), 523–531. doi: 10.1016/0092-8674(81)90394-9

Kingston, R. E., Nierman, W. C., and Chamberlin, M. J. (1981). A direct effect of guanosine tetraphosphate on pausing of Escherichia coli RNA polymerase during RNA chain elongation. *J. Biol. Chem.* 256 (6), 2787–2797.

Lacour, S., and Landini, P. (2004). SigmaS-dependent gene expression at the onset of stationary phase in Escherichia coli: function of sigmaS-dependent genes and identification of their promoter sequences. *J. Bacteriol.* 186 (21), 7186–7195. doi: 10.1128/JB.186.21.7186-7195.2004

Lalaouna, D., Eyraud, A., Devinck, A., Prevost, K., and Masse, E. (2019). GcvB small RNA uses two distinct seed regions to regulate an extensive targetome. *Mol. Microbiol.* 111 (2), 473–486. doi: 10.1111/mmi.14168

Landini, P., Egli, T., Wolf, J., and Lacour, S. (2014). sigmaS, a major player in the response to environmental stresses in Escherichia coli: role, regulation and mechanisms of promoter recognition. *Environ. Microbiol. Rep.* 6 (1), 1–13. doi: 10.1111/1758-2229.12112

Lange, R., and Hengge-Aronis, R. (1994). The cellular concentration of the sigma S subunit of RNA polymerase in Escherichia coli is controlled at the levels of transcription, translation, and protein stability. *Genes Dev.* 8 (13), 1600–1612. doi: 10.1101/gad.8.13.1600

Leavitt, R. I., and Umbarger, H. E. (1962). Isoleucine and valine metabolism in Escherichia coli. XI. Valine inhibition of the growth of Escherichia coli strain K-12. *J. Bacteriol.* 83, 624–630. doi: 10.1128/JB.83.3.624-630.1962

Lee, H. J., and Gottesman, S. (2016). sRNA roles in regulating transcriptional regulators: Lrp and SoxS regulation by sRNAs. *Nucleic Acids Res.* 44 (14), 6907–6923. doi: 10.1093/nar/gkw358

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25 (14), 1754–1760. doi: 10.1093/bioinformatics/btp324

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25 (4), 402–408. doi: 10.1006/meth.2001.1262

Maaløe, O. (1979). "Regulation of the protein synthesizing machinery - ribosomes, tRNA, factors and so on.," in *Biological regulation and development*. Ed. R. F. Goldberger. (New York: Plenum Publishing Corp), 487–542.

Maruyama, H., and Mizuno, D. (1970). Ribosome degradation and the degradation products in starved Escherichia coli. I. Comparison of the degradation rate and of the nucleotide pool between Escherichia coli B and Q-13 strains in phosphate deficiency. *Biochim. Biophys. Acta* 199 (1), 159–165. doi: 10.1016/0005-2787(70)90704-5

MATLAB (Release R2016b). *MATLAB and Statistics Toolbox*, The MathWorks, Inc., Natick, Massachusetts, United States.

Motorin, Y., Muller, S., Behm-Ansmant, I., and Branlant, C. (2007). Identification of modified residues in RNAs by reverse transcription-based methods. *Methods Enzymol.* 425, 21–53. doi: 10.1016/S0076-6879(07)25002-5

Murray, K. D., and Bremer, H. (1996). Control of spoT-dependent ppGpp synthesis and degradation in *Escherichia coli*. *J. Mol. Biol.* 259 (1), 41–57. doi: 10.1006/jmbi.1996.0300

Neidhardt, F. C., Bloch, P. L., and Smith, D. F. (1974). Culture medium for enterobacteria. *J. Bacteriol.* 119 (3), 736–747. doi: 10.1128/JB.119.3.736-747.1974

Nomura, M., Gourse, R., and Baughman, G. (1984). Regulation of the synthesis of ribosomes and ribosomal components. *Annu. Rev. Biochem.* 53 (1), 75–117. doi: 10.1146/annurev.bi.53.070184.000451

Nyström, T., and Neidhardt, F. C. (1992). Cloning, mapping and nucleotide sequencing of a gene encoding a universal stress protein in Eschericha coli. *Mol. Microbiol.* 6 (21), 3187–3198. doi: 10.1111/j.1365-2958.1992.tb01774.x

Paley, S., Parker, K., Spaulding, A., Tomb, J.-F., O'Maille, P., and Karp, P. D. (2017). The Omics Dashboard for interactive exploration of gene-expression data. *Nucleic Acids Res.* 45 (21), 12113–12124. doi: 10.1093/nar/gkx910

Piir, K., Paier, A., Liiv, A., Tenson, T., and Maivali, U. (2011). Ribosome degradation in growing bacteria. *EMBO Rep.* 12 (5), 458–462. doi: 10.1038/embor.2011.47

Roghanian, M., Zenkin, N., and Yuzenkova, Y. (2015). Bacterial global regulators DksA/ppGpp increase fidelity of transcription. *Nucleic Acids Res.* 43 (3), 1529–1536. doi: 10.1093/nar/gkv003

Ross, W., Sanchez-Vazquez, P., Chen, A. Y., Lee, J. H., Burgos, H. L., and Gourse, R. L. (2016). ppGpp binding to a site at the RNAP-DksA interface accounts for its dramatic effects on transcription initiation during the stringent response. *Mol. Cell* 62 (6), 811–823. doi: 10.1016/j.molcel.2016.04.029

Ryals, J., Little, R., and Bremer, H. (1982). Control of rRNA and tRNA syntheses in *Escherichia coli* by guanosine tetraphosphate. *J. Bacteriol.* 151 (3), 1261–1268. doi: 10.1128/JB.151.3.1261-1268.1982

Sørensen, M. A., Jensen, K. F., and Pedersen, S. (1994). High concentrations of ppGpp decrease the RNA chain growth rate: Implications for protein synthesis and translation fidelity during amino acid starvation in *Escherichia coli*. *J. Mol. Biol.* 236, 441–454. doi: 10.1006/jmbi.1994.1156

Sanchez-Vazquez, P., Dewey, C. N., Kitten, N., Ross, W., and Gourse, R. L. (2019). Genome-wide effects on Escherichia coli transcription from ppGpp binding to its two sites on RNA polymerase. *Proc. Natl. Acad. Sci.* 116, 8310–8319, 201819682. doi: 10.1073/pnas.1819682116

Sands, M. K., and Roberts, R. B. (1952). The effects of a tryptophan-histidine deficiency in a mutant of Escherichia coli. *J. Bacteriol.* 63 (4), 505–511. doi: 10.1128/JB.63.4.505-511.1952

Sarubbi, E., Rudd, K. E., and Cashel, M. (1988). Basal ppGpp level adjustment shown by new spoT mutants affect steady state growth rates and rrnA ribosomal promoter regulation in Escherichia coli. *Mol. Gen. Genet. MGG* 213 (2), 214–222. doi: 10.1007/bf00339584

Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270 (5235), 467–470. doi: 10.1126/science.270.5235.467

Stenum, T. S., Sorensen, M. A., and Svenningsen, S. L. (2017). Quantification of the abundance and charging levels of transfer RNAs in Escherichia coli. *J. Vis. Exp.* (126). doi: 10.3791/56212

Svenningsen, S. L., Kongstad, M., Stenum, T. S., Munoz-Gomez, A. J., and Sorensen, M. A. (2017). Transfer RNA is highly unstable during early amino acid starvation in Escherichia coli. *Nucleic Acids Res*. 45, 793–804 doi: 10.1093/nar/gkw1169

Tian, C., Roghanian, M., Jørgensen, M. G., Sneppen, K., Sørensen, M. A., Gerdes, K., et al. (2016). Rapid curtailing of the stringent response by toxin-antitoxin encoded mRNases. *J. Bacteriol.* 198, 1918–1926. doi: 10.1128/jb.00062-16

Traxler, M. F., Summers, S. M., Nguyen, H. T., Zacharia, V. M., Hightower, G. A., Smith, J. T., et al. (2008). The global, ppGpp-mediated stringent response to amino acid starvation in Escherichia coli. *Mol. Microbiol.* 68 (5), 1128–1148. doi: 10.1111/j.1365-2958.2008.06229.x

Traxler, M. F., Zacharia, V. M., Marquardt, S., Summers, S. M., Nguyen, H. T., Stark, S. E., et al. (2011). Discretely calibrated regulatory loops controlled by ppGpp partition gene induction across the 'feast to famine' gradient in Escherichia coli. *Mol. Microbiol.* 79 (4), 830–845. doi: 10.1111/j.1365-2958.2010.07498.x

Tschowri, N., Lindenberg, S., and Hengge, R. (2012). Molecular function and potential evolution of the biofilm-modulating blue light-signalling pathway of Escherichia coli. *Mol. Microbiol.* 85 (5), 893–906. doi: 10.1111/j.1365-2958.2012.08147.x

Valle, J., Da Re, S., Schmid, S., Skurnik, D., D'Ari, R., and Ghigo, J. M. (2008). The amino acid valine is secreted in continuous-flow bacterial biofilms. *J. Bacteriol.* 190 (1), 264–274. doi: 10.1128/JB.01405-07

Vogel, U., and Jensen, K. F. (1994). Effects of guanosine 3',5'-bisdiphosphate (ppGpp) on rate of transcription elongation in isoleucine-starved Escherichia coli. *J. Biol. Chem.* 269 (23), 16236–16241.

Vollmer, A. C., and Bark, S. J. (2018). "Chapter One - Twenty-Five Years of Investigating the Universal Stress Protein: Function, Structure, and Applications," in *Advances in Applied Microbiology*. Eds. S. Sariaslani and G. M. Gadd. (Cambridge, MA, United States: Academic Press), 1–36.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10 (1), 57–63. doi: 10.1038/nrg2484

Weber, H., Polen, T., Heuveling, J., Wendisch, V. F., and Hengge, R. (2005). Genome-wide analysis of the general stress response network in Escherichia coli: sigmaS-

dependent genes, promoters, and sigma factor selectivity. *J. Bacteriol.* 187 (5), 1591–1603. doi: 10.1128/JB.187.5.1591-1603.2005

Wild, J., Hennig, J., Lobocka, M., Walczak, W., and Klopotowski, T. (1985). Identification of the dadX gene coding for the predominant isozyme of alanine racemase in Escherichia coli K12. *Mol. Gen. Genet.* 198 (2), 315–322. doi: 10.1007/bf00383013

Williams, G. J., Breazeale, S. D., Raetz, C. R., and Naismith, J. H. (2005). Structure and function of both domains of ArnA, a dual function decarboxylase and a formyltransferase, involved in 4-amino-4-deoxy-L-arabinose biosynthesis. *J. Biol. Chem.* 280 (24), 23000–23008. doi: 10.1074/jbc.M501534200

Winther, K. S., Roghanian, M., and Gerdes, K. (2018). Activation of the Stringent Response by Loading of RelA-tRNA Complexes at the Ribosomal A-Site. *Mol. Cell* 95–105 (1), e104, 95–105. doi: 10.1016/j.molcel.2018.02.033

Wren, D. J. D., and Conway, T. (2006). Meta-analysis of published transcriptional and translational fold changes reveals a preference for low-fold inductions. *OMICS: A J. Integr. Biol.* 10 (1), 15–27. doi: 10.1089/omi.2006.10.15

Zundel, M. A., Basturea, G. N., and Deutscher, M. P. (2009). Initiation of ribosome degradation during starvation in Escherichia coli. *RNA* 15 (5), 977–983. doi: 10.1261/rna.1381309

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

REVIEW
published: 31 March 2020
doi: 10.3389/fgene.2020.00226

Translation-Targeting RiPPs and Where to Find Them

Dmitrii Y. Travin[1,2], Dmitry Bikmetov[3,4] and Konstantin Severinov[1,3,4,5]*

[1] Center of Life Sciences, Skolkovo Institute of Science and Technology, Moscow, Russia, [2] Institute of Gene Biology, Russian Academy of Sciences, Moscow, Russia, [3] Institute of Molecular Genetics, Russian Academy of Sciences, Moscow, Russia, [4] Center for Precision Genome Editing and Genetic Technologies for Biomedicine, Institute of Gene Biology, Russian Academy of Sciences, Moscow, Russia, [5] Waksman Institute for Microbiology, Rutgers, Piscataway, NJ, United States

Prokaryotic translation is among the major targets of diverse natural products with antibacterial activity including several classes of clinically relevant antibiotics. In this review, we summarize the information about the structure, biosynthesis, and modes of action of translation inhibiting ribosomally synthesized and post-translationally modified peptides (RiPPs). Azol(in)e-containing RiPPs are known to target translation, and several new compounds inhibiting the ribosome have been characterized recently. We performed a systematic search for biosynthetic gene clusters (BGCs) of azol(in)e-containing RiPPs. This search uncovered several groups of clusters that likely direct the synthesis of novel compounds, some of which may be targeting the ribosome.

Keywords: RiPPs, ribosome, antibiotics, LAPs, YcaO, azol(in)e-modified peptides, genome mining

OPEN ACCESS

Edited by:
Michael Ibba,
The Ohio State University,
United States

Reviewed by:
Matthieu Gagnon,
The University of Texas Medical Branch at Galveston, United States
Jesko Köhnke,
Helmholtz-Institute for Pharmaceutical Research Saarland (HIPS), Germany

****Correspondence:***
Dmitrii Y. Travin
dmitrii.travin@skoltech.ru

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *16 November 2019*
Accepted: *26 February 2020*
Published: *31 March 2020*

Citation:
Travin DY, Bikmetov D and Severinov K (2020) Translation-Targeting RiPPs and Where to Find Them. Front. Genet. 11:226. doi: 10.3389/fgene.2020.00226

INTRODUCTION

Antibiotics are extensively used worldwide in healthcare, agriculture, and food preservation. However, development and the spread of resistance to most antibiotics discovered during the second half of the 20th century in the course of the so-called "golden era of antibiotics" has become a global threat (Brown and Wright, 2016). With multinational pharmaceutical corporations exiting the field, the search for novel natural products, which remain a major source of novel bioactive compounds including antibiotics (Li and Vederas, 2009; Moloney, 2016), is largely concentrated in academia and small companies (Wright, 2017). Classical activity-based strain screening approaches, which are costly and which often result in a rediscovery of the already known compounds, are giving way to "smarter" techniques (Baltz, 2019). The genome mining approach relies on a "from genes to products" paradigm, which is the opposite of conventional activity-based antibiotic searches, and critically depends on the rapid accumulation of genomic data in publicly available databases (Ziemert et al., 2016).

Genome mining for novel metabolites begins with *in silico* predictions of functions of the groups of genes called "biosynthetic gene clusters" (BGCs), whose products may take part in the biosynthesis of a certain metabolite. Making specific predictions about the structure of the final compound enables better prioritizing of candidate BGCs for subsequent time-consuming downstream experimental validation. Both proteinogenic and non-proteinogenic amino acids can act as building blocks for the production of specialized metabolites, leading to a great diversity of naturally occurring bioactive peptides. Peptide natural products originating from bacteria and fungi and currently used as antibiotics are dominated by non-ribosomal peptides (NRPs), assembled by large multisubunit enzymatic complexes (Süssmuth and Mainz, 2017). With the number of known NRP BGCs steadily growing, our abilities to predict both the amino acid sequence and tailoring modifications of final compounds based on the sequences of NRP synthases and additional enzymes

encoded in BGCs improve as well; however, the complete "nonribosomal code" is not yet known (Ackerley et al., 2016).

In addition to NRPs, ribosomally synthesized and post-translationally modified peptides (RiPPs) comprise another rapidly expanding class of bioactive peptides. They are produced by posttranslational modifications (PTMs) of ribosomally synthesized precursors by dedicated enzyme machinery (Arnison et al., 2013). Compared to those of NRPs, RiPP BGCs are generally smaller and contain a precursor peptide gene, which enables better prediction of the final product structure based not only on the properties of enzymes involved in biosynthesis but also on the chemical structure of the initial peptide substrate. While the number of identified RiPPs grows, there is still an enormous space for the discovery of new compounds expanding the diversity within the already known RiPP subclasses and of entirely new groups of RiPPs harboring novel modifications, as evidenced, for example, by the recently described ranthipeptides (Hudson et al., 2019) and streptide-like RiPPs (Schramma et al., 2015).

Among the RiPPs exhibiting antibacterial activity, there are examples of compounds targeting various validated drug targets, including bacterial RNA polymerase (microcin J25; Delgado et al., 2001), DNA gyrase [microcin B17 (McB); Heddle et al., 2001], and the cell membrane (lanthibiotics; Chatterjee et al., 2005). The prokaryotic ribosome, another key target of many antibiotics currently in use (reviewed in Wilson, 2014; Polikanov et al., 2018), is also inhibited by certain RiPPs. In addition to RiPPs directly interacting with the ribosome, there are those that block translation by binding to elongation factors or inhibiting the activity of aminoacyl-tRNA synthetases. In this review, we briefly summarize the available data on the structure, biosynthesis, mode of action, and BGC composition of the known RiPPs inhibiting different steps of translation (**Figure 1**). In addition, we specifically explore the genomic landscape of azol(in)e-containing RiPPs in an attempt to predict novel RiPPs inhibiting translation. Although definitive predictions of the compound mode of action based entirely on the genomic data can be made only in rare cases (see, for example, the self-resistance guided identification of new topoisomerase inhibitors; Panter et al., 2018), the results of our search define new subclasses of azol(in)e-containing RiPPs, which may include the translation inhibitors.

Thiopeptides

Thiopeptides comprise one of the best-studied subclasses of RiPPs, with more than 100 compounds characterized to date. Produced predominantly by *Actinobacteria*, they demonstrate various activities including antibacterial and antiplasmoidal (Gaillard et al., 2016), which result from their ability to inhibit translation by the prokaryotic ribosome and the ribosome within the apicoplast of the malaria parasite *Plasmodium falciparum* (Clough et al., 1997). All known thiopeptides have the specific set of biosynthetic genes in their BGCs (**Figure 2A**), share common structural features (**Figure 2B**), and use two major mechanisms for translation inhibition: they either interact directly with the ribosome (**Figure 2C**) or prevent the binding of aminoacyl-tRNA by the elongation factor EF-Tu (**Figure 2D**).

The first group of thiopeptides is characterized by the small size of a macrocycle (26 atoms) and a conserved region (**Figure 2B**, nosiheptide, red dashed frame) essential for the interaction with their binding site on the large ribosome subunit, GTP-ase associated center (GAC). Thiostrepton, nosiheptide, and micrococcin, all belonging to this group, were co-crystallized with the large ribosome subunit of *Deinococcus radiodurans*. While most ribosome-targeting antibiotics contact only rRNA, these thiopeptides interact with both rRNA and ribosomal proteins, binding in a cleft formed by the N-terminal domain of the ribosomal protein uL11 and the loops of helices H43 and H44 of the 23S rRNA (**Figure 2C**; Harms et al., 2008). Their binding site overlaps with the binding sites of IF2, EF-G, and EF-Tu. Consistently, thiopeptides inhibit initiation (Brandi et al., 2004), translocation (Rodnina et al., 1999), and tRNA delivery to the ribosome (Gonzalez et al., 2007). The practical applications of naturally occurring thiopeptides are limited by their poor water solubility (Just-Baringo et al., 2014). Nevertheless, nosiheptide is used as a growth stimulating additive in mixed animal food, while thiostrepton is used to treat skin infections in animals.

GE2270A, isolated in the early 1990s from actinobacterium *Planobispora rosea* (Selva et al., 1991), is a representative of the second functional group of thiopeptides. It binds to the elongation factor EF-Tu in a complex with GTP and prevents the formation of the ternary complex with aminoacyl-tRNA (**Figure 2D**; Heffron and Jurnak, 2000). A derivative of GE2770A with an altered C-terminus named LFF571, proved to be effective and safe in phase II clinical trials against *Clostridium difficile* infections, a rare case when a RiPP-inspired molecule reached clinical trials (Jarrad et al., 2015). The GE2270A binding site is located between EF-Tu domains I and III and partially overlaps with the binding site of polyketide antibiotic pulvomycin, an interesting example of two chemically unrelated compounds adopting a similar mode of inhibition of the same molecular target (Parmeggiani et al., 2006). Other thiopeptides demonstrating similar modes of action, e.g., thiomuracin and GE37468A, are all characterized by the medium size of the macrocycle (29 atoms) and the presence of a conserved Asn or MeAsn residue (**Figure 2B**, GE2770A, red dashed frame), required for the interaction with EF-Tu (Young et al., 2012).

The "core" set of PTMs characterizing most thiopeptides includes the installation of azol(in)e heterocycles, dehydration of amino acids, and macrocyclization *via* the formation of a six-membered azacycle. Azole cycles (**Figure 2B**, red) are synthesized in a two-step reaction from amino acids with a nucleophilic group in their side chains. First, an YcaO-domain cyclodehydratase together with ThiF-like partner protein (the latter is required for precursor recognition) converts Cys residues into thiazolines and Ser or Thr residues into oxazolines or methyloxazolines, respectively. Flavin mononucleotide (FMN)-dependent dehydrogenase can further oxidize azoline cycles into aromatic azoles. The mechanism of azole installation and diversity of azol(in)e-containing RiPPs were recently reviewed by Burkhart et al. (2017).

The formation of dehydrated amino acids (dehydroalanine from Ser residues and dehydrobutyrine from Thr, **Figure 2B**, pink) in thiopeptides proceeds *via* a glutamylation-elimination

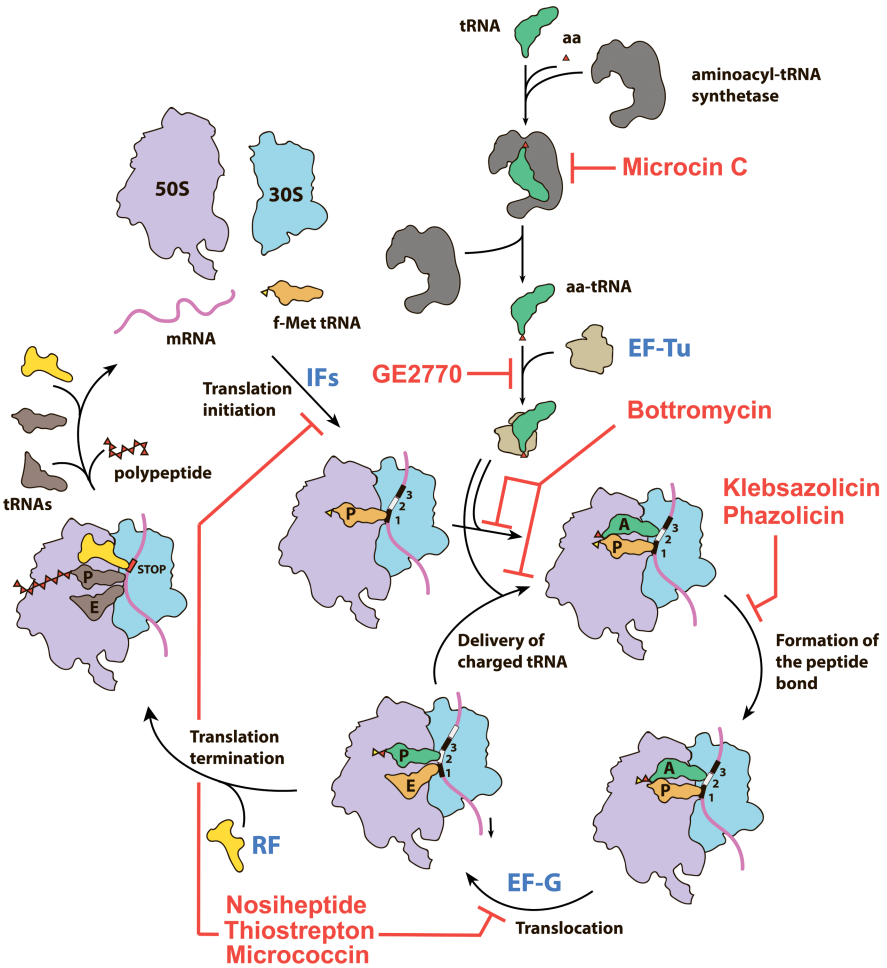

FIGURE 1 | Overview of the prokaryotic translation process and its steps targeted by various RiPPs. IFs—initiation factors, RF—release factor, aa—amino acid, aa-tRNA—aminoacyl tRNA, f-Met tRNA—initiator N-formylmethionine tRNA.

mechanism with $tRNA^{Glu}$ functioning as a donor of glutamyl (Hudson et al., 2015). The enzymes catalyzing this reaction are encoded by two separate genes (so-called "split" LanB), which are also found in BGCs of other, unrelated, RiPPs including class I lanthipeptides. The mechanisms and enzymology of dehydration of amino acids were reviewed by Repka et al. (2017). Amino acid dehydration is a prerequisite for the most remarkable modification of thiopeptides—formation of the central nitrogen-containing heterocycle (**Figure 2B**, green). This reaction follows the [4+2] cycloaddition mechanism (aza-Diels-Alder reaction) that is in most cases accompanied by the removal of the leader peptide and leads to the formation of a macrocycle system. The enzymes responsible for catalysis of these reactions in various biosynthetic pathways were reviewed by Jeon et al. (2017). The structure of the central azacycle is the basis for thiopeptide classification into series (from *a* to *e*) (Bagley et al., 2005).

The genes encoding the enzymes responsible for "core" PTMs mentioned above, together with a precursor peptide gene, constitute the simplest variant of thiopeptide BGC, i.e., the *laz*-cluster of lactazole synthesis (**Figure 2A**, Hayashi et al., 2014). Most thiopeptide BGCs are larger and encode the enzymes catalyzing additional modifications (**Figure 2A**, *nos*- and *tbd*-clusters) as well as transporters and regulatory proteins. The tailoring modifications may include the formation of a side ring system *via* the addition of indole derivates (**Figure 2B**, blue), modifications of the C-terminus to prevent hydrolysis by carboxypeptidases, glycosylation, hydroxylation, etc. The diversity and mechanisms of thiopeptide PTMs were reviewed in detail by Zheng et al. (2017).

Linear Azol(in)e-Containing Peptides (LAPs)

The name "linear azol(in)e-containing peptides (LAPs)" refers to the only two characteristics shared by compounds from this diverse subgroup of RiPPs: they (*i*) have azol(in)e cycles installed along the polypeptide backbone and (*ii*) do not undergo macrocyclization (Arnison et al., 2013). Thus, a minimal LAP BGC comprises only a gene encoding the precursor peptide

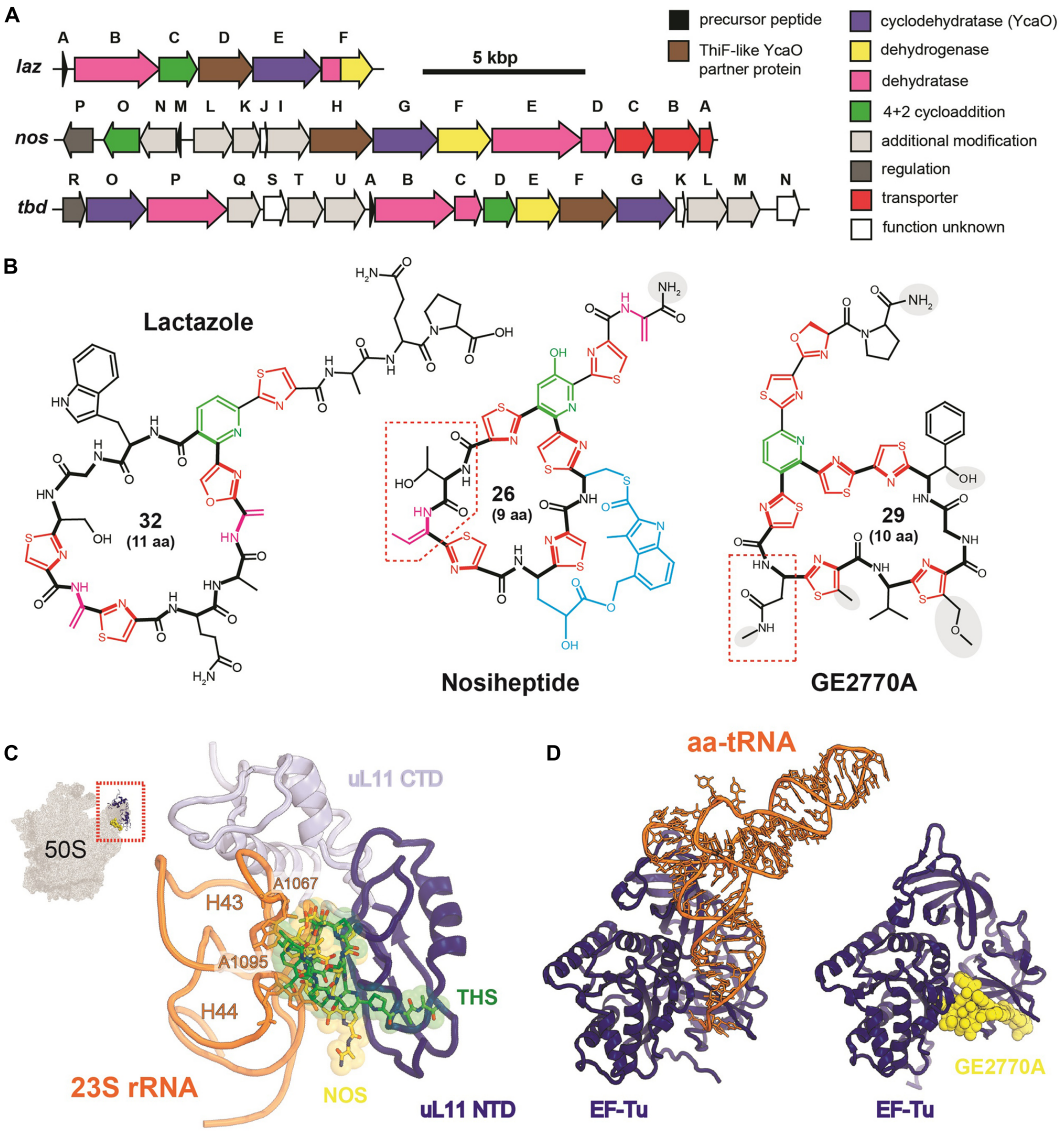

FIGURE 2 | Thiopeptides. **(A)** Biosynthetic gene clusters of lactazole (a minimal thiopeptide-encoding BGC), nosiheptide, and GE2270A. Functions of encoded proteins are listed on the right. **(B)** Chemical structures of lactazole, nosiheptide, and GE2270A. Azol(in)e cycles are shown in red, six-membered central azacycles in green, dehydrated amino acids in pink, and methylindole acid-containing second ring system of nosiheptide in blue. Auxiliary tailoring modifications are highlighted with gray background. The number of atoms and amino acid residues in the macrocyclic system is indicated for each compound. Macrocycles are shown in bold. Red dashed polygons show conserved residues characteristic for ribosome targeting (nosiheptide) and EF-Tu targeting (GE2270A) thiopeptides. **(C)** Mode of nosiheptide and thiostrepton interaction with the ribosome (PDB IDs 2ZJP and 3CF5; Harms et al., 2008). Nosiheptide is yellow, thiostrepton is green, uL11 ribosomal protein CTD (C-terminal domain) is lightblue, NTD (N-terminal domain)—blue, H43 and H44 helices of 23S rRNA are orange, residues A1067 and A1095 (*E. coli* nomenculature), involved in the rRNA-antibiotic interaction are shown as sticks. **(D)** Mode of action of GE2770A. Elongation factor EF-Tu (blue) is shown in complex with aminoacyl tRNA (aa-tRNA, orange) and GE2270A (yellow) (PDB IDs 1B23 and 2C77, respectively, Nissen et al., 1999; Parmeggiani et al., 2006). The binding of GE2270A prevents the interaction of EF-Tu with aa-tRNA acceptor stem.

(gene A) and gene(s) coding for the enzymes involved in the installation of azole cycles: a YcaO-cyclodehydratase (the product of the D gene), which in most cases has a partner protein that is required for leader peptide recognition (either an E1-like protein, the product of the C gene, or a ThiF-like protein encoded by the F gene), and a dehydrogenase (the product of the B gene) which oxidizes azolines to azoles. In some BGCs, genes coding for C and D proteins, are fused and code for a single polypeptide (Burkhart et al., 2017). The list of additional modifications of LAPs is diverse and includes, among others, N-methylation (plantazolicin, Lee et al., 2013), the formation of dehydroamino acids and N-terminal acetylation (goadsporin, Ozaki et al., 2016), N-terminal oxyme formation, and C-terminal O-methylation (azolemycin, Liu et al., 2016).

As the set of chemical characteristics required to attribute a compound to LAPs is not particularly restrictive, the group

includes compounds without any obvious sequence similarity of peptide precursors. The relationships between the known LAPs (to date there are less than two dozens of well-characterized compounds) resemble a "sea with islands," where each "island" is formed by a group of the closely related homologs (e.g., streptolysin S with its relatives; Molloy et al., 2011) without links between the "islands."

Klebsazolicin (KLB) is the first characterized translation-targeting LAP (Metelev et al., 2017). Its BGC was found in the genome of *Klebsiella pneumonia* sub. *ozaenae* and contains a gene for the precursor peptide (*klpA*), genes encoding the enzymes required for azole cycle installation (*klpCBD*), and an exporter pump gene (*klpE*) (**Figure 3A**). In addition to three thiazoles and one oxazole cycle, KLB has an N-terminal amidine cycle formed by the first two residues of the core part of the precursor peptide (Ser1 and Gln2; **Figure 3B**), a modification unique among the known LAPs. *In vitro* studies have demonstrated that this cycle is formed after the proteolytic cleavage of the leader peptide and strictly requires the YcaO-domain KlpD cyclodehydratase (Travin et al., 2018). The amidine cycle is required for KLB to function since the derivatives with a full set of azole cycles, but lacking the amidine cycle, do not inhibit translation. As it is typical for other LAPs, KLB is a narrow spectrum antibiotic: it is active against the genera closely related to *Klebsiella*, including *Yersinia* and *Escherichia*.

Cocrystallization of KLB with the *Thermus thermophilus* ribosome (which because of ease of crystallization is widely used for structural studies of ribosome-targeting compounds) revealed the molecular details of its mode of action. KLB binds in the upper part of the peptide exit tunnel in a site adjacent to the peptidyl-transferase center (PTC) (**Figure 3C**). Acting as a cork in the bottle, KLB blocks the passage of the nascent peptide, only allowing the synthesis of di- or tripeptides that remain associated with tRNA and stay bound to the elongating ribosome.

Phazolicin (PHZ) is another recently discovered ribosome-targeting LAP produced by soil bacterium *Rhizobium* sp. Pop5, a symbiont of wild beans *Phaseolus vulgaris* (Travin et al., 2019). In terms of the overall composition, PHZ BGC is identical to that of KLB (**Figure 3A**). PHZ is a 27-amino acid long peptide, every third amino acid of which is converted into an azole cycle. Unlike KLB, no modifications other than Cys and Ser side chain cyclizations are present in PHZ (**Figure 3B**). PHZ is active against various rhizobia that are closely related to the producing strain. Similarly, to KLB, PHZ targets the ribosome exit tunnel but does this through a different set of interactions, which were revealed by cryo-EM of the *Escherichia coli* ribosome complex with PHZ (**Figure 3C**). Four azole cycles of PHZ form a π–π stacking system, which stabilizes 3D globular structure of the peptide, while the three azoles are involved in stacking with nucleobases of the 23S rRNA. Unlike KLB, PHZ has three positively charged residues involved in the interactions with phosphates and other polar groups of 23S rRNA. PHZ also interacts with the loop regions of two ribosomal proteins (uL4 and uL22). Amino acid sequences of these loops confer the species-specific mode of translation inhibition by PHZ, which, unlike KLB, does not bind to *T. thermophilus* ribosome.

Bottromycins

Bottromycins are extensively modified RiPPs that exhibit potent antimicrobial activity against the drug-resistant human pathogens including vancomycin-resistant *Enterococcus* (VRE) and methicillin-resistant *Staphylococcus aureus* (MRSA) (Shimamura et al., 2009). In early works, bottromycin A2 was demonstrated to inhibit protein synthesis both *in vitro* and *in vivo* (Tanaka et al., 1966). Further studies showed that the action of bottromycin does not interfere with the peptide bond formation and translocation steps. Bottromycins are believed to bind in the A-site of the ribosome (Otaka and Kaji, 1976, 1981, 1983) and block the interaction of aminoacyl-tRNAs with the ribosome, almost an unexploited target among the currently used antibiotics. However, further structural studies of bottromycin mechanism of action are needed to establish the details of this interaction at a molecular level, as previous studies used indirect approaches sometimes leading to contradictory conclusions.

Although the first representative of the bottromycin family of RiPPs was isolated from *Streptomyces bottropensis* in 1957 (Waisvisz et al., 1957), more than 50 years passed until the structure of the compound was finally confirmed by total chemical synthesis (Shimamura et al., 2009). Bottromycins are eight-amino acid long extensively modified peptides originating from the N-terminal part of a precursor peptide (thus bottromycin precursor has a "follower" peptide to which modification machinery binds, rather than N-terminal "leader" common among other RiPPs). The biosynthesis of bottromycin includes many steps and was intensively studied using the untargeted methabolomics approach (Crone et al., 2016) and *in vitro* reconstitution of separate modification reactions. The PTMs characteristic to bottromycins include the formation of the N-terminal macroamidine cycle (**Figure 4A**, green) and C-terminal thiazole (**Figure 4A**, red) catalyzed by two divergent YcaO-domain enzymes acting without any partner proteins (so-called "standalone YcaOs") (Franz et al., 2017; Schwalen et al., 2017). In addition to these cyclizations, C_β-methylations of Pro, Phe, and Val residues, as well as O-methylation of aspartate take place (**Figure 4A**, gray background; Huo et al., 2012). Different methylation profiles lead to multiple forms of bottromycins produced by the same strain (Eyles et al., 2018). In addition to the genes encoding YcaO heterocyclases and methyltransferases, bottromycin BGC includes genes encoding an enzyme, which removes the N-terminal methionine residue (Mann et al., 2016), an amidohydrolase required for the follower peptide removal (Sikandar et al., 2019), a cytochrome performing oxidative decarboxylation of the C-terminal azoline into azole, and a transporter (**Figure 4A**).

Microcin C and Related Compounds

Microcin C (McC) is a peptide-nucleotide antibiotic produced by *E. coli* strains bearing a plasmid with a six-gene *mcc* gene cluster (**Figure 4B**), which encodes a seven amino acid-long precursor peptide (MccA, MRTGNAN), enzymes responsible for its PTM (MccB, MccD, and MccE), an exporter pump (MccC), and a peptidase providing autoimmunity (MccF). The product of the *mccA* gene is adenylated by MccB, which leads to the formation

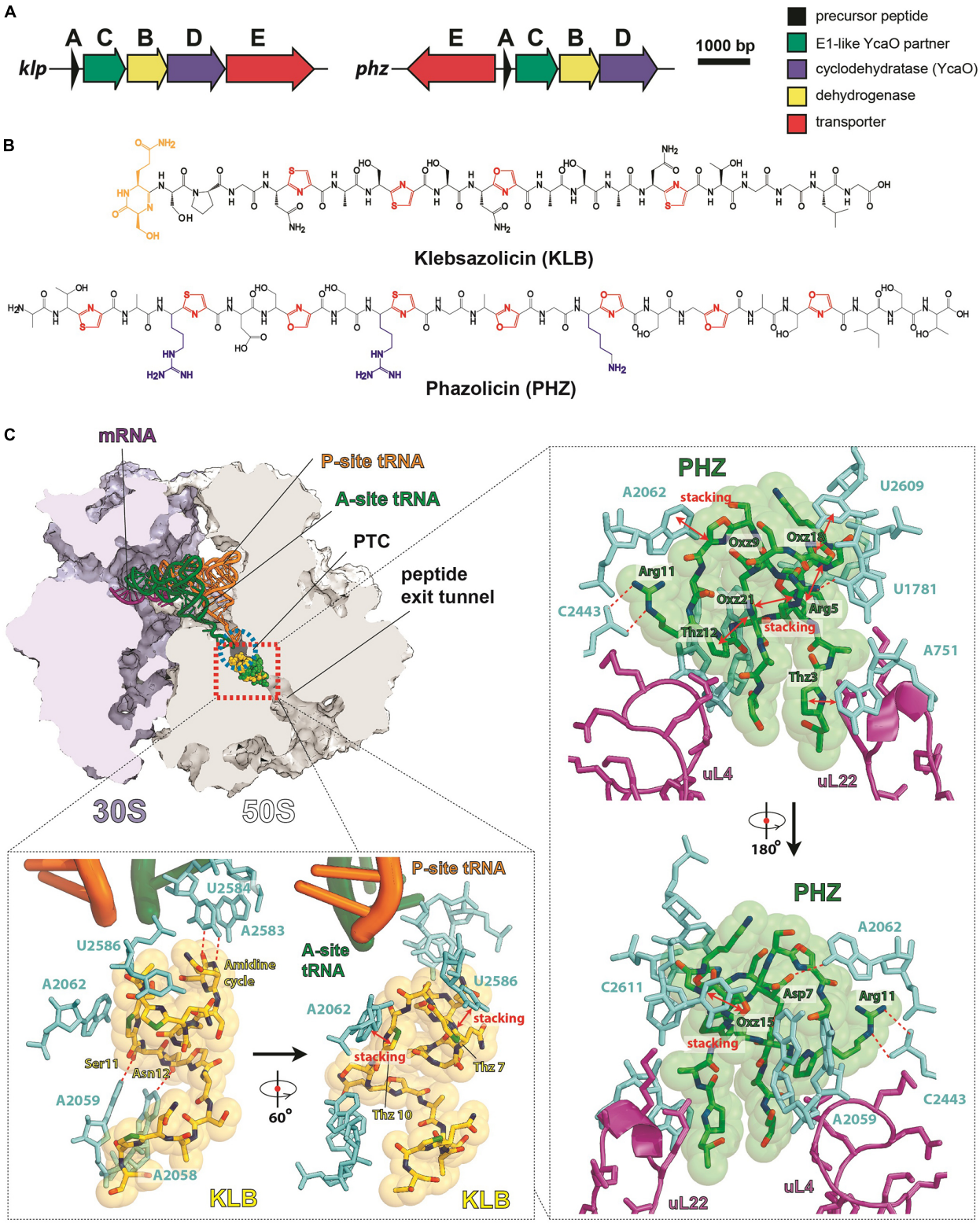

FIGURE 3 | LAPs targeting the ribosome exit tunnel. **(A)** BGCs of klebsazolicin (*klpACBDE*) and phazolicin (*phzEACBD*). The functions of encoded proteins are listed on the right. **(B)** Chemical structures of klebsazolicin (KLB) and phazolicin (PHZ), azole cycles are shown in red, positively charged amino acid side chains of PHZ—in blue, the N-terminal amidine cycle of KLB is yellow. **(C)** Mechanism of translation inhibition by KLB (PDB ID 5W4K, Metelev et al., 2017) and PHZ (PDB ID 6U48, Travin et al., 2019). Interactions of the two antibiotics with the ribosome are shown: π–π stacking is denoted by red arrows and hydrogen bonds with red dashed lines. Nucleobases of 23S rRNA are cyan, ribosomal proteins are magenta, and tRNAs are orange and green. PTC—peptidyl-transferase center.

of a non-hydrolyzable N-P bond between C-terminal asparagine and phosphate (Roush et al., 2008). MccD and MccE are required for additional decoration of the molecule with aminopropyl group attached to the phosphate (Kulikovsky et al., 2014). Recent studies increased the number of McC-related compounds: RiPPs of this family undergoing cytidylation instead of adenylation were discovered, and carboxymethylation of the cytidine was shown to be an additional tailoring step required for an optimal bioactivity (Serebryakova et al., 2016; Tsibulskaya et al., 2017).

Microcin C is a Trojan-horse antibiotic imported into sensitive cells *via* the inner membrane transporter YejABEF, which recognizes the peptide part of McC (Novikova et al., 2007). The McC molecule itself is not toxic for the cell; the peptide part has to be deformylated and subsequently degraded by non-specific cellular oligopeptidases (Kazakov et al., 2008) to release a nonhydrolyzable analog of aspartyl adenylate, a potent inhibitor of aspartyl-tRNA synthetase (**Figure 4B**, gray background) (Metlitskaya et al., 2006). This leads to the accumulation of

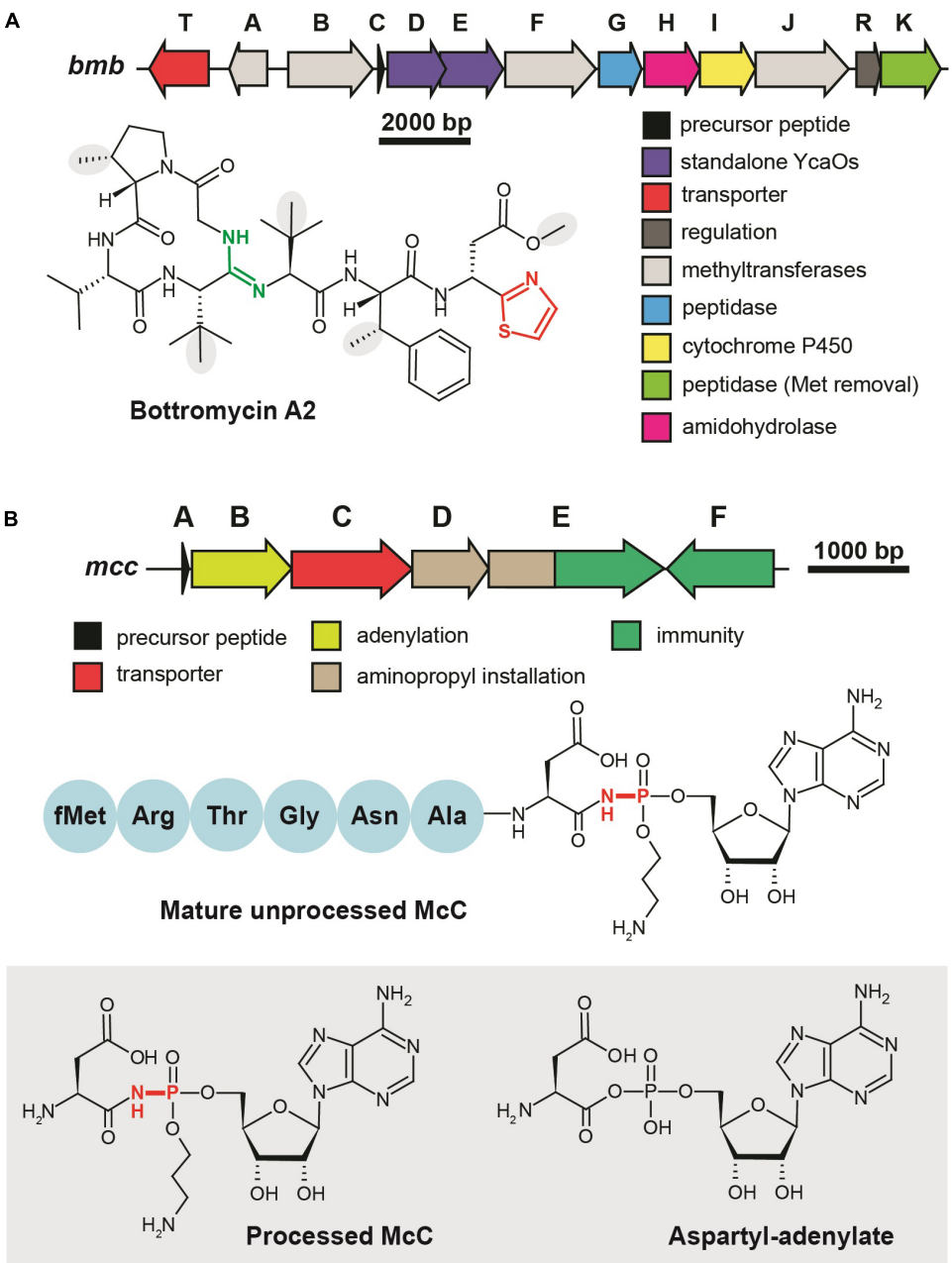

FIGURE 4 | Translation-targeting RiPPs without structural information on target binding **(A)** BGC and chemical structure of bottromycin A2 from *Streptomyces bottropensis*. C-terminal thiazole is shown in red, macroamidine bond is green, and C_β-methyl groups are shown on the gray background. **(B)** BGC of microcin C, chemical structures of unprocessed microcin C and of processed form, an analog of aspartyl adenylate. The non-hydrolyzable N–P bond is shown in red.

uncharged $tRNA^{Asp}$, inhibition of protein synthesis, and the cessation of cell growth. Thus, McC is another example of a RiPP (together with GE2270A discussed earlier), which does not directly interact with the ribosome but blocks translation by inhibiting the supply of substrates required for protein synthesis. Although McC has been studied for more than 30 years and structures of multiple enzymes involved in its biosynthesis and immunity have been determined (Agarwal et al., 2011, 2012; Dong et al., 2019), we still lack structural information about the details of McC interaction with aspartyl-tRNA synthetase.

GENOME MINING FOR NOVEL TRANSLATION INHIBITING RIPPS

In light of the data discussed above, it is evident that many translation-targeting RiPPs contain azol(in)e cycles. At least in several cases where the mode of interaction with a ribosome is known, these cycles take part in stacking interactions with rRNA nucleobases thus mediating binding of the inhibitor to the target. We decided to perform a search for novel groups of azol(in)e-containing RiPPs in the genomes present in publicly available databases with a goal of identifying the putative translation inhibitors as well as other bioactive molecules. Due to their essential role in azol(in)e-containing RiPP biosynthesis, the genes encoding the YcaO-domain-containing enzymes were chosen as a starting point for our search.

In less than 10 years YcaO-domain containing enzymes went from being DUFs (domains of unknown function) to one of the most studied groups of RiPP modification proteins (Burkhart et al., 2017). It was demonstrated that YcaO enzymes play the key role in the catalysis of three distinct reactions of PTM of proteins and peptides including the installation of azoline cycles (Dunbar et al., 2012), amidines (Burkhart et al., 2017; Franz et al., 2017; Travin et al., 2018), and thioamides (Mahanta et al., 2018; Schwalen et al., 2018). A common mechanism involving the nucleophilic attack on the amide bond containing substrate with a subsequent ATP-dependent phosphorylation of the intermediate followed by phosphate elimination underlies all these activities. Three groups of proteins are regarded as YcaO partners, allowing for the interaction of the enzyme with its substrate (the recognition of the leader peptide in case of RiPP biosynthesis). These are E1-like proteins and ThiF-like proteins, fused or clustered together with azoline-forming YcaOs (Burkhart et al., 2017), and TfuA-like proteins considered to be a hallmark of the BGCs of thioamidated compounds (Santos-Aberturas et al., 2019).

To identify new BGCs of azol(in)e-containing compounds, we started with a sensitive search for sequences of YcaO domain-containing enzymes present in genomes from the RefSeq database (O'Leary et al., 2016). In brief, the subsequent steps included filtering, clusterization, and annotation of genomic regions surrounding the recovered *ycaO* genes (**Figure 5**). To visualize the obtained diversity and to identify families of BGCs, we constructed a sequence similarity network of all YcaO-containing BGCs, which was then analyzed manually (for detailed description of procedures, see section "Methods"). A curated set of characterized YcaO-containing BGCs including those present in the MIBiG database (Kautsar et al., 2019) or described elsewhere in the literature (including previous bioinformatic predictions) was used as a reference (**Supplementary Table S1**). In the current study, we focused only on clusters containing E1-like or a ThiF-like partner proteins and did not consider TfuA-containing BGCs or BGCs with standalone YcaOs. We also did not consider BGCs of thiopeptides and closely related RiPPs (defined as clusters containing *lanB*-like genes) as they were recently searched with various tools (Li et al., 2012; Schwalen et al., 2018). The genomic landscape of all azol(in)e-containing peptides was studied by Cox et al. (2015), however, since the time of this publication new azol(in)e-containing RiPPs with characterized modes of action (including ribosome-targeting KLB and PHZ) have been discovered, and many more sequenced genomes have been deposited in publicly available databases. Moreover, several improved methods and software have become available.

Figure 6 represents a similarity network of YcaO-containing BGCs encoding E1-like (**Figure 6A**) or ThiF-like (**Figure 6B**) YcaO partner proteins (see **Supplementary Table S2** for the list of all BGCs). BGCs of the already characterized compounds from the curated dataset are shown as blue circles. In the network with E1-like partners, these include BGCs of bioactive antibacterials McB, KLB, and PHZ as well as a number of streptolysin S-like RiPPs (clostridilysin S; listeriolysin S; Cotter et al., 2008; Gonzalez et al., 2010) and hakacin, whose biosynthesis was studied *in vitro*, but the structure of the naturally produced compound remains unknown (Melby et al., 2012; Dunbar and Mitchell, 2013). In

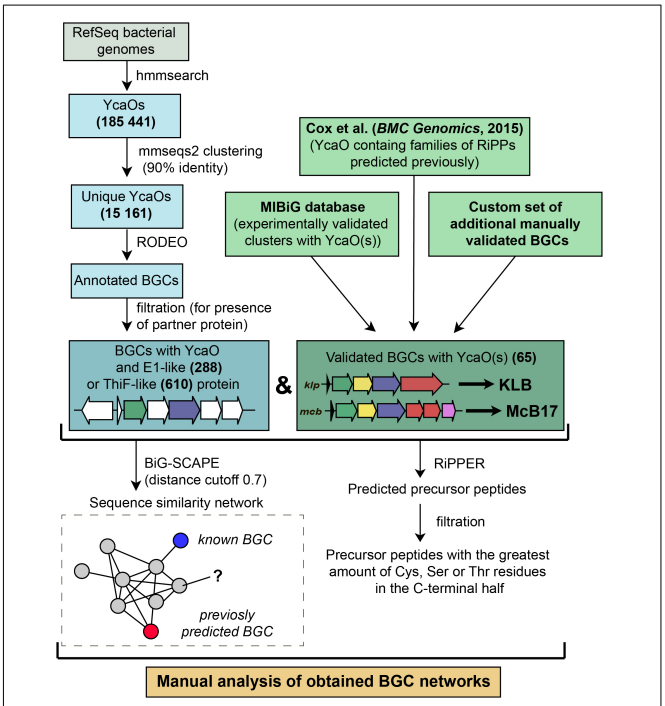

FIGURE 5 | A workflow for identification of YcaO-containing RiPP BGCs.

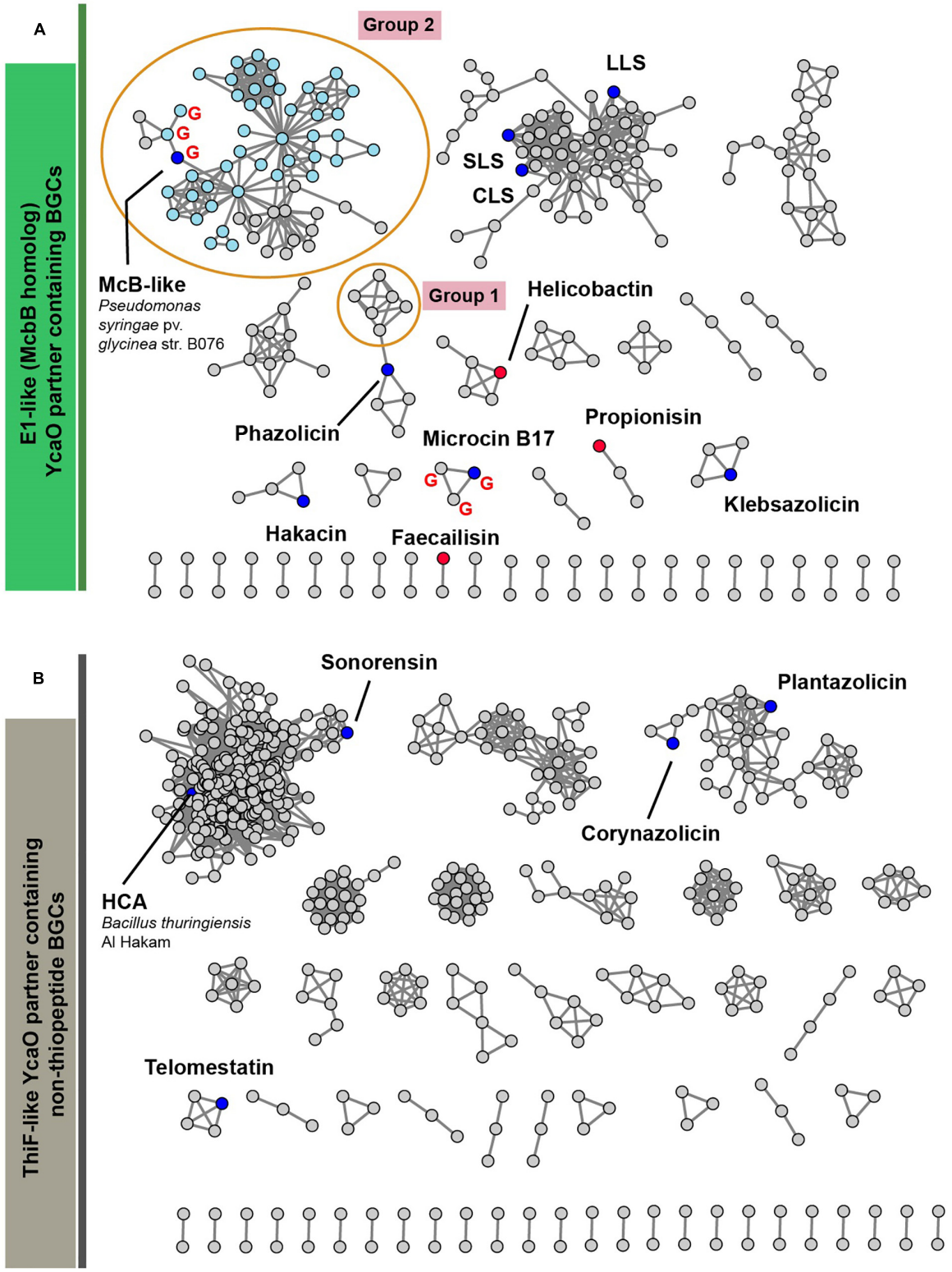

FIGURE 6 | A similarity network of YcaO-containing BGCs with E1-like **(A)** and ThiF-like **(B)** partner proteins. Nodes representing BGCs of already characterized compounds are shown in blue, nodes representing BGCs analyzed in Cox et al. (2015) are red. BGCs containing an *mcbG*-homolog are denoted with red "G." Groups of clusters discussed in the text are shown in orange ellipses. Light blue color of nodes in group 2 shows BGCs from genus *Pseudomonas*. SLS—strepolysin S, LLS—listeriolysin S, CLS—clostridiolysin S, HCA—heterocycloanthracin.

the network of ThiF-like protein containing BGCs we observed a large group of heterocycloanthracin (HCA) BGCs, which include the already characterized sonorensin (Chopra et al., 2014) and HCA from *Bacillus thuringiensis* Al Hakam (Dunbar et al., 2015).

Below, we discuss three groups of BGCs, which attracted our attention during the analysis of the networks and putative peptides encoded by these BGCs as predicted by RiPPER (Santos-Aberturas et al., 2019) in **Figure 6**. We consider it likely that the first group of these BGCs encodes new translation targeting RiPPs; the second may also do so, while the third was so interesting in terms of RiPP encoding clusters' evolution, that we could not help but discuss it in this article.

Lactazolicins

The first group of BGCs contains clusters from the representatives of the genus *Lactobacillus*, which form a connected component with PHZ BGC (**Figure 6**, Group 1). Analysis of these BGCs and their homologs from genera *Enterococcus* and *Streptococcus* found with an additional BLAST search revealed that all these BGCs share the same set of genes, which, in addition to modification machinery and export pump homologs of those in PHZ BGC (**Figure 7A**, genes E, C, B, and D_2), includes three auxiliary genes (**Figure 7A**, genes X_1, D_1, and X_2). The product of gene D_1 is the second YcaO protein. It is distinct from the product of the D_2 gene and lacks the C-terminal PxP-motif, found in azoline-forming YcaOs and involved in catalysis (Ghilarov et al., 2019). According to the results of HHPred (Söding et al., 2005), the product of gene X_2 is distantly related to ThiF/MccB/PaaA proteins and contains a RiPP recognition element (RRE) — a domain found in different RiPP modification enzymes binding leader peptides (Burkhart et al., 2015). The presence of the second YcaO and of the X_2 gene product, which could function either as a partner protein or an independent adenylating enzyme (Ghodge et al., 2016; Dong et al., 2019), makes additional modifications of the precursor peptide highly probable. We were unable to detect any homologs of the X_1 gene product among the known proteins.

Following the conventional practice of giving names to the proposed new groups of compounds (Cox et al., 2015) and in accordance with the nomenclature recommended for LAPs (Arnison et al., 2013), we named this group of putative translation inhibitors *lactazolicins*. All lactazolicin clusters encode 83-106 amino acid-long putative precursor peptides with 8–12 repeats of the [Cxxx] motif in the N-terminal part of the predicted core segment (**Figure 7A**). HCAs represent an already known group of RiPPs, which have a similar pattern of repeated cysteine residues in the core part (Haft, 2009). However, HCA precursors (also found in our search, **Figure 6B**, the largest group of BGCs) have the [Cxx] motif repeated rather than [Cxxx], and the overall composition of HCA BGCs also differs significantly from that of lactazolicin BGCs. Unlike HCAs, where the [Cxx]-repeat containing part of the precursor is rich in glycines, the N-terminal [Cxxx] repeat-containing part of lactazolicin precursors is enriched in positively charged amino acids (Arg, Lys). In the cases of PHZ and proline-rich peptides (which do not belong to RiPPs but also target the ribosome exit tunnel) (Gagnon et al., 2016), the side chains of positively charged amino acids take part in the interaction with phosphate groups of rRNA. We thus hypothesize that lactazolicins may also affect translation.

Microcin B17-Like BGCs From Pseudomonads

Microcin B17 is a DNA-gyrase-targeting LAP produced by some strains of *E. coli*. The McB BGC contains a set of enzymes similar to those encoded by the KLB and PHZ BGCs and an additional gene *mcbG*, which encodes a pentapeptide repeat protein (PRP) (Li et al., 1996; Heddle et al., 2001). McbG is likely a DNA mimic that decreases the formation of toxic gyrase-DNA complexes trapped by McB, thus protecting the gyrase in the McB-producing cell (Hegde et al., 2005; Vetting et al., 2011). Clusters similar to that of McB were described in the genomes of several pathovars of *Pseudomonas syringae* and their products also target gyrase (Metelev et al., 2013).

A relatively large network of clusters retrieved by our search (**Figure 6A**, Group 2) contains no previously characterized representatives except for an *mcb*-operon homolog from *P. syringae* (blue circle). However, several of these clusters (marked with red letter G) contain a gene coding for a PRP protein. The overall sequence similarity and the distribution of potentially cyclizable residues in precursor peptides from clusters with and without the PRP gene differ (**Figure 7B**). Thus, it is highly probable that *mcb*-like clusters without a PRP gene encode a RiPP with a target distinct from DNA gyrase. While we cannot establish whether these are translation-targeting RiPPs, compounds with the same set of proteins in their BGC (KLB and PHZ) do affect translation.

Flavazolicins

The last group of putative new LAP BGCs was identified during the analysis of precursor peptides predicted with RiPPER (Santos-Aberturas et al., 2019). The precursor peptide identified in the genome of flavobacterium *Algibacter aquaticus* SK-16 (a singlet and therefore not shown in **Figure 6A**; **Figure 8A**) appears to have resulted from a duplication of a standard leader-core ancestral precursor gene (**Figure 8B**). As a result, in a single ORF, there are two putative core sequences rich in Ser and Cys residues separated by an "internal" leader (another leader is N-terminally located) (**Figure 8C**). A similar cassette-like arrangement of core peptides has been described for several different groups of RiPPs including cyanobactins (Gu et al., 2018), thiovarsolines (Santos-Aberturas et al., 2019), orbitides (Shim et al., 2015), and dikaritins (Ding et al., 2016); but in all these cases, precursors are composed of a single leader, followed by several core peptides, interspersed by signal sequences required for the cleavage of each core at C- and N-termini by dedicated peptidases (**Figure 8B** shows, as an example, the sequence of TruE1 — the precursor of patellins 2 and 3, representatives of cyanobactins).

A BLAST search for similar BGCs resulted in identifying six additional BGCs, that share the same set of modification enzymes (**Figure 8A**). The first three originate from the genomes of *Flavobacteriaceae* closely related to *Algibacter*, while three others were found in the genomes of *Gammaproteobacteria*. Interestingly, only two of these clusters contained a fused

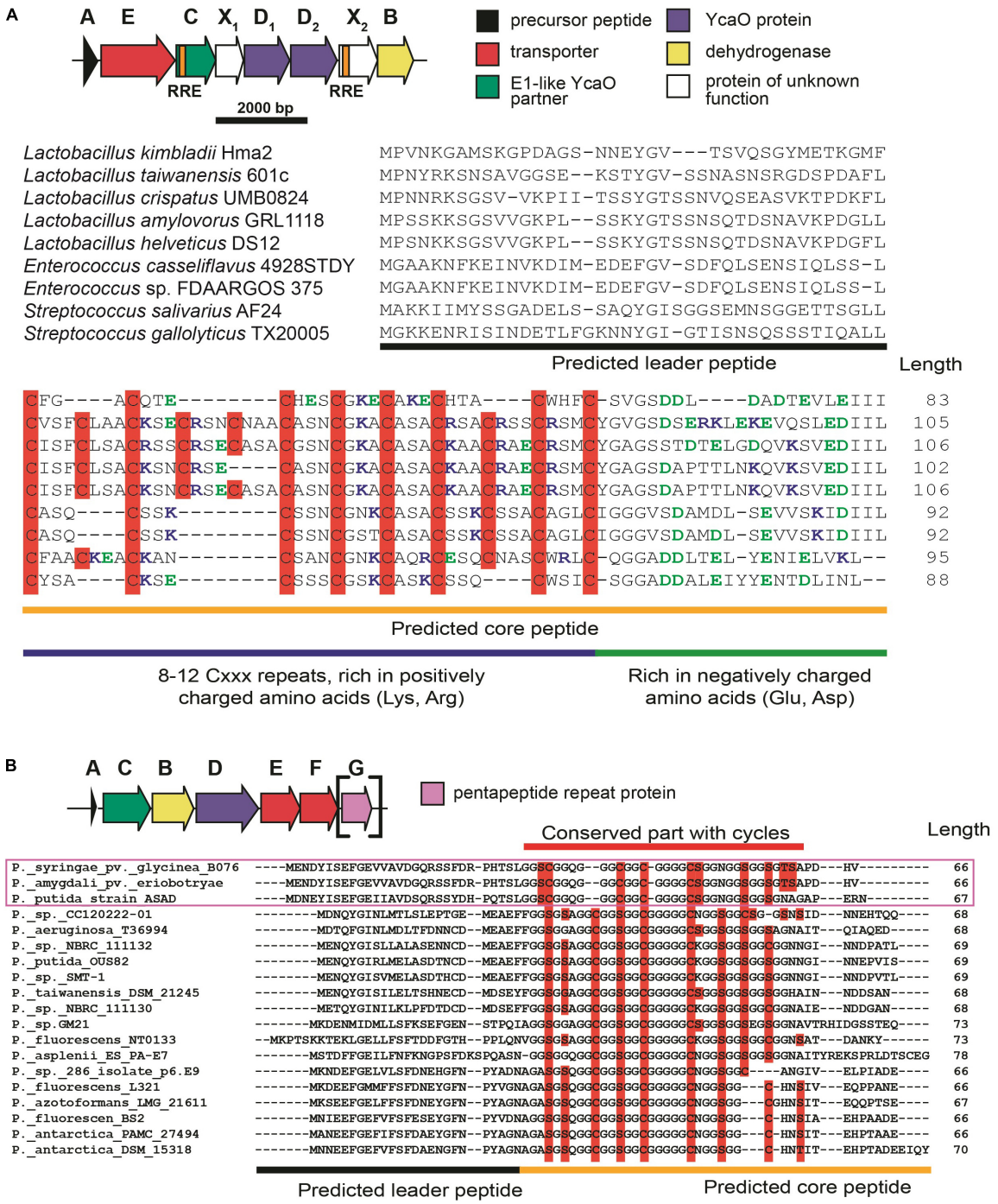

FIGURE 7 | (A) Common BGC composition conserved among lactazolicin BGCs, proposed functions of the encoded proteins are listed on the right. RiPP recognition elements (RREs) in C and X_2 genes are depicted. Alignment of precursor peptides of lactazolicins, predicted leader and core parts are shown. Cysteines in the core part are shown on red background, positively charged amino acids of the predicted core are blue, and negatively charged residues are green. **(B)** The composition of McB-like BGC from pseudomonads. *mcbG* homolog is shown in square brackets as it is not present in the majority of clusters from pseudomonads, the genes are colored according to the color scheme in **(A)**. Alignment of precursor peptides predicted with RiPPER, those encoded in PRP gene-containing BGCs are in magenta frame. Potentially cyclized residues are shown with red background, predicted core and leader parts are shown.

precursor peptide gene, while the rest had a set of one to three separate ORFs encoding non-fused precursor peptides (**Figure 8B**). These different genome arrangements from the closely related species provide a glimpse on how the genes of cassette-containing peptides may originate from an independent single short ORF through gene duplication

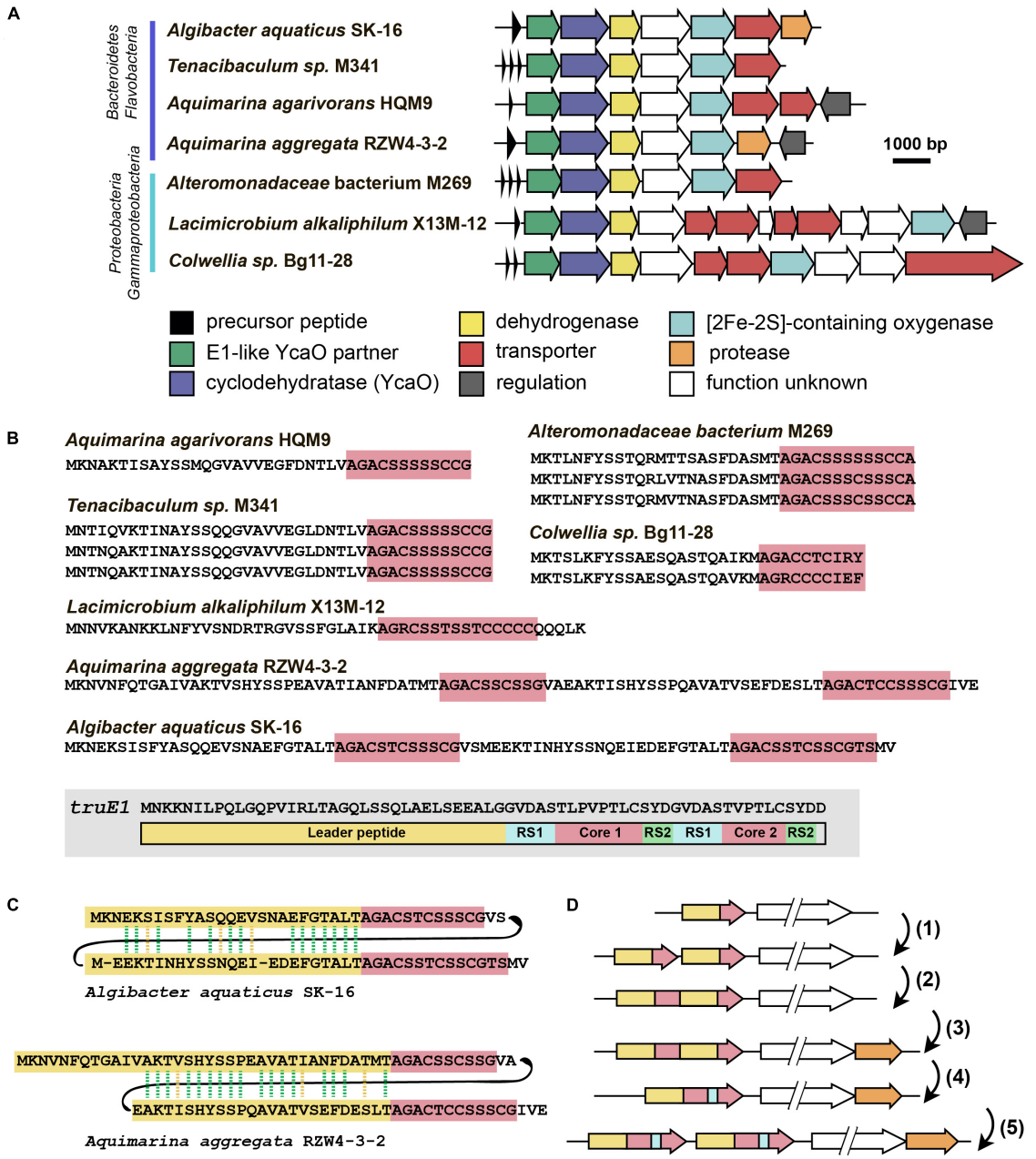

FIGURE 8 | Flavazolicines. **(A)** Comparison of biosynthetic gene clusters encoding a putative new group of LAPs found in *Flavobacteria* and *Gammaproteobacteria* genomes. Predicted functions of the Encoded proteins are listed below. **(B)** Comparison of precursor peptides of flavazolicins. Conserved core sequences containing cyclizable residues are shown with red background. The precursor peptide sequence of cyanobactins patellins 2 and 3 (*truE1* gene product) is shown for comparison on the gray background. Functional parts of the peptide including leader (yellow), two cores (red), and recognition sequences of peptidases (RS1 and RS2) are shown (Gu et al., 2018). **(C)** Sequences of cassette-containing precursor peptides of flavazolicins showing the conserved positions in two leader sequences. Conserved positions are shown with green dashed lines, synonymous substitutions with yellow dashed lines. **(D)** A possible scenario in the evolution of cassette-containing peptides. See main text for the explanations.

[**Figure 8D**(1)], fusion [**Figure 8D**(2)], and a subsequent reduction of the role of the internal leader to that of a recognition sequence of proteases [**Figure 8D**(4)]. Further multiplication of cassette-containing precursor genes may lead to arrangements found in several cyanobactin clusters [**Figure 8D**(5); Gu et al., 2018].

Strikingly, only the BGCs with fused precursors contain an additional gene, which is a predicted protease (**Figure 8A**, orange). This enzyme may be involved in the processing required to produce individual modified core parts. The acquisition of an additional protease gene may be the step that follows the fusion of two independent ORFs in the course of cassette-containing BGC

evolution [**Figure 8D**(3)]. We named the products of this family of BGCs *flavazolicins*. Characterizing the products encoded in these BGCs and establishing the details of their biosynthesis and function appears to be an exciting direction of future work.

CONCLUDING REMARKS

Although the number of the known subclasses and unique representatives of RiPPs increases each year, a remarkable proportion of publications devoted to novel compounds provides information only about the structure and sometimes evaluates the bioactivity of a modified peptide. Researchers focusing on RiPP clusters as a source of unprecedented enzymatic activities rarely proceed toward establishing the mode of action of the target compound and are even less likely to establish its physiological or ecological role. Addressing these questions is a challenging task, which partially explains the lack of detailed information about the precise mechanisms of action for many groups of RiPPs, including some that are known and have been studied for decades (e.g., bottromycin, McB). We hope that the upcoming years will provide more structural insights not only on the enzymology of RiPP modification widely studied now, but also on the principles the already known and novel compounds act by.

In many cases, the analysis of genomic information was a starting point for further successful discoveries of a novel RiPP, facilitating the prediction of the BGC product based on the sequences of precursor peptides and modification enzymes. Through genome mining, future studies will not only result in the discovery of new compounds but will also allow systemization of our knowledge about RiPP genomic landscape and a better understanding of RiPP clusters' evolutionary relations.

METHODS

Search for YcaO Containing BGCs, Filtration, and Annotation

146,381 bacterial genomes were downloaded from RefSeq (O'Leary et al., 2016) database on 27 March 2019. To obtain all YcaO domain-containing proteins we searched the database with profile HMMs (TIGR03549, TIGR03604, and PF02624) from public databases using hmmer package[1]. We clustered resulting hits with mmseqs2 (Mirdita et al., 2019) (90% identity; 90% coverage) to remove duplicates and redundant highly similar sequences from organisms, which genome sequences are overrepresented in the database.

The genomic regions of 12.5 kbp to each side of the identified unique YcaO protein-coding genes were annotated with RODEO (Tietz et al., 2017) using Pfam 32.0 and TIGRFAMs 15.0 databases. For further analysis, we selected genomic regions according to several rules. First, we collected regions that encode proteins containing E1-like (PF00881, TIGR03603, TIGR04424) or ThiF-like (PF00899, TIGR02354, TIGR02356, TIGR03693, TIGR03736, TIGR03882) domains. Initial search was very sensitive and false positive results were obtained. Thus, we removed predicted YcaO proteins that were not annotated with TIGR03549, TIGR03604, or PF02624 domains in the RODEO output. In order to exclude thiopeptides, studied comprehensively in several other works, we removed genomic regions containing genes of lantipeptide dehydratase (PF14028, PF04738, TIGR03897, PF05147). Putative precursor peptides were predicted with RiPPER (Santos-Aberturas et al., 2019). For each BGC, the best predicted precursor peptide was selected as the one bearing the highest number of cyclizable residues (Ser, Thr, Cys residues) within the C-terminal half.

Using a custom script (available on http://github.com/bikdm12/RODEO2antiSMASH) we converted RODEO output to genbank files imitating antiSMASH (Blin et al., 2019) output. The script adds a feature "cluster" with information about the class of the product. The coordinates of this feature are boundaries of the group of genes located on the same strand not farther than 100 bps from each other and containing YcaO protein. Also, genes that may be related to azol(in)e-containing RiPP biosynthesis (for the list of domains see **Supplementary Table S3**) were marked as biosynthetic. These files were then used to build a sequence similarity network with BiG-SCAPE (Navarro-Muñoz et al., 2020) subsequently visualized with Cytoscape (Shannon et al., 2003).

[1]hmmer.org

AUTHOR CONTRIBUTIONS

DT and DB performed the bioinformatic search and interpreted its results. DT prepared the figures. All authors wrote the manuscript.

FUNDING

This work was supported by RSF grant number 19-14-00266 to Svetlana Dubiley.

ACKNOWLEDGMENTS

We thank Dr. Dmitry Ghilarov, Ms. Elena Smertina, and Dr. Svetlana Dubiley for critical reading of the manuscript and valuable suggestions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00226/full#supplementary-material

TABLE S1 | Manually curated set of previously characterized YcaO-containing BGCs (including bioinformatically predicted, but not validated experimentally).

TABLE S2 | List of YcaO-containing BGCs with E1-like or ThiF-like YcaO partner proteins identified in this study (exluding clusters encoding lantipeptide dehydratases).

TABLE S3 | List of Pfam and TIGRFAMs profile HMMs used to mark biosynthetic genes.

REFERENCES

Ackerley, D. F. (2016). Cracking the nonribosomal code. *Cell Chem. Biol.* 23, 535–537. doi: 10.1016/j.chembiol.2016.05.001

Agarwal, V., Metlitskaya, A., Severinov, K., and Nair, S. K. (2011). Structural basis for microcin C7 inactivation by the MccE acetyltransferase. *J. Biol. Chem.* 286, 21295–21303. doi: 10.1074/jbc.M111.226282

Agarwal, V., Tikhonov, A., Metlitskaya, A., Severinov, K., and Nair, S. K. (2012). Structure and function of a serine carboxypeptidase adapted for degradation of the protein synthesis antibiotic microcin C7. *Proc. Natl. Acad. Sci. U.S.A.* 109, 4425–4430. doi: 10.1073/pnas.1114224109

Arnison, P. G., Bibb, M. J., Bierbaum, G., Bowers, A. A., Bugni, T. S., Bulaj, G., et al. (2013). Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat. Prod. Rep.* 30, 108–160. doi: 10.1039/c2np20085f

Bagley, M. C., Dale, J. W., Merritt, E. A., and Xiong, X. (2005). Thiopeptide antibiotics. *Chem. Rev.* 105, 685–714. doi: 10.1021/cr0300441

Baltz, R. H. (2019). Natural product drug discovery in the genomic era: realities, conjectures, misconceptions, and opportunities. *J. Ind. Microbiol. Biotechnol.* 46, 281–299. doi: 10.1007/s10295-018-2115-4

Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., et al. (2019). antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* 47, W81–W87. doi: 10.1093/nar/gkz310

Brandi, L., Marzi, S., Fabbretti, A., Fleischer, C., Hill, W. E., Gualerzi, C. O., et al. (2004). The translation initiation functions of IF2: targets for thiostrepton inhibition. *J. Mol. Biol.* 335, 881–894. doi: 10.1016/j.jmb.2003.10.067

Brown, E. D., and Wright, G. D. (2016). Antibacterial drug discovery in the resistance era. *Nature* 529, 336–343. doi: 10.1038/nature17042

Burkhart, B. J., Hudson, G. A., Dunbar, K. L., and Mitchell, D. A. (2015). A prevalent peptide-binding domain guides ribosomal natural product biosynthesis. *Nat. Chem. Biol.* 11, 564–570. doi: 10.1038/nchembio.1856

Burkhart, B. J., Schwalen, C. J., Mann, G., Naismith, J. H., and Mitchell, D. A. (2017). YcaO-dependent posttranslational amide activation: biosynthesis. *Struct. Funct. Chem. Rev.* 117, 5389–5456. doi: 10.1021/acs.chemrev.6b00623

Chatterjee, C., Paul, M., Xie, L., and van der Donk, W. A. (2005). Biosynthesis and mode of action of lantibiotics. *Chem. Rev.* 105, 633–684. doi: 10.1021/cr030105v

Chopra, L., Singh, G., Choudhary, V., and Sahoo, D. K. (2014). Sonorensin: an antimicrobial peptide, belonging to the heterocycloanthracin subfamily of bacteriocins, from a new marine isolate, Bacillus sonorensis MT93. *Appl. Environ. Microbiol.* 80, 2981–2990. doi: 10.1128/AEM.04259-13

Clough, B., Strath, M., Preiser, P., Denny, P., and Wilson, I. R. (1997). Thiostrepton binds to malarial plastid rRNA. *FEBS Lett.* 406, 123–125. doi: 10.1016/s0014-5793(97)00241-x

Cotter, P. D., Draper, L. A., Lawton, E. M., Daly, K. M., Groeger, D. S., Casey, P. G., et al. (2008). Listeriolysin S, a novel peptide haemolysin associated with a subset of lineage I *Listeria monocytogenes*. *PLoS Pathog.* 4:e1000144. doi: 10.1371/journal.ppat.1000144

Cox, C. L., Doroghazi, J. R., and Mitchell, D. A. (2015). The genomic landscape of ribosomal peptides containing thiazole and oxazole heterocycles. *BMC Genomics* 16:778. doi: 10.1186/s12864-015-2008-0

Crone, W. J. K., Vior, N. M., Santos-Aberturas, J., Schmitz, L. G., Leeper, F. J., and Truman, A. W. (2016). Dissecting bottromycin biosynthesis using comparative untargeted metabolomics. *Angew. Chem. Int. Ed.* 55, 9639–9643. doi: 10.1002/anie.201604304

Delgado, M. A., Rintoul, M. R., Farías, R. N., and Salomón, R. A. (2001). *Escherichia coli* RNA polymerase is the target of the cyclopeptide antibiotic microcin J25. *J. Bacteriol.* 183, 4543–4550. doi: 10.1128/JB.183.15.4543-4550.2001

Ding, W., Liu, W.-Q., Jia, Y., Li, Y., van der Donk, W. A., and Zhang, Q. (2016). Biosynthetic investigation of phomopsins reveals a widespread pathway for ribosomal natural products in Ascomycetes. *Proc. Natl. Acad. Sci. U.S.A.* 113, 3521–3526. doi: 10.1073/pnas.1522907113

Dong, S.-H., Kulikovsky, A., Zukher, I., Estrada, P., Dubiley, S., Severinov, K., et al. (2019). Biosynthesis of the RiPP trojan horse nucleotide antibiotic microcin C is directed by the N-formyl of the peptide precursor. *Chem. Sci.* 10, 2391–2395. doi: 10.1039/c8sc03173h

Dunbar, K. L., Melby, J. O., and Mitchell, D. A. (2012). YcaO domains use ATP to activate amide backbones during peptide cyclodehydrations. *Nat. Chem. Biol.* 8, 569–575. doi: 10.1038/nchembio.944

Dunbar, K. L., and Mitchell, D. A. (2013). Insights into the mechanism of peptide cyclodehydrations achieved through the chemoenzymatic generation of amide derivatives. *J. Am. Chem. Soc.* 135, 8692–8701. doi: 10.1021/ja4029507

Dunbar, K. L., Tietz, J. I., Cox, C. L., Burkhart, B. J., and Mitchell, D. A. (2015). Identification of an auxiliary leader peptide-binding protein required for azoline formation in ribosomal natural products. *J. Am. Chem. Soc.* 137, 7672–7677. doi: 10.1021/jacs.5b04682

Eyles, T. H., Vior, N. M., and Truman, A. W. (2018). Rapid and robust yeast-mediated pathway refactoring generates multiple new bottromycin-related metabolites. *ACS Synth. Biol.* 7, 1211–1218. doi: 10.1021/acssynbio.8b00038

Franz, L., Adam, S., Santos-Aberturas, J., Truman, A. W., and Koehnke, J. (2017). Macroamidine formation in bottromycins is catalyzed by a divergent YcaO enzyme. *J. Am. Chem. Soc.* 139, 18158–18161. doi: 10.1021/jacs.7b09898

Gagnon, M. G., Roy, R. N., Lomakin, I. B., Florin, T., Mankin, A. S., and Steitz, T. A. (2016). Structures of proline-rich peptides bound to the ribosome reveal a common mechanism of protein synthesis inhibition. *Nucleic Acids Res.* 44, 2439–2450. doi: 10.1093/nar/gkw018

Gaillard, T., Madamet, M., Tsombeng, F. F., Dormoi, J., and Pradines, B. (2016). Antibiotics in malaria therapy: which antibiotics except tetracyclines and macrolides may be used against malaria? *Malar. J.* 15:556. doi: 10.1186/s12936-016-1613-y

Ghilarov, D., Stevenson, C. E. M., Travin, D. Y., Piskunova, J., Serebryakova, M., Maxwell, A., et al. (2019). Architecture of microcin B17 synthetase: an octameric protein complex converting a ribosomally synthesized peptide into a DNA gyrase poison. *Mol. Cell* 73, 749–762. doi: 10.1016/j.molcel.2018.11.032

Ghodge, S. V., Biernat, K. A., Bassett, S. J., Redinbo, M. R., and Bowers, A. A. (2016). Post-translational claisen condensation and decarboxylation en route to the bicyclic core of pantocin A. *J. Am. Chem. Soc.* 138, 5487–5490. doi: 10.1021/jacs.5b13529

Gonzalez, D. J., Lee, S. W., Hensler, M. E., Markley, A. L., Dahesh, S., Mitchell, D. A., et al. (2010). Clostridiolysin S, a post-translationally modified biotoxin from Clostridium botulinum. *J. Biol. Chem.* 285, 28220–28228. doi: 10.1074/jbc.M110.118554

Gonzalez, R. L., Chu, S., and Puglisi, J. D. (2007). Thiostrepton inhibition of tRNA delivery to the ribosome. *RNA* 13, 2091–2097. doi: 10.1261/rna.499407

Gu, W., Sardar, D., Pierce, E., and Schmidt, E. W. (2018). Roads to rome: role of multiple cassettes in cyanobactin ripp biosynthesis. *J. Am. Chem. Soc.* 140, 16213–16221. doi: 10.1021/jacs.8b09328

Haft, D. H. (2009). A strain-variable bacteriocin in *Bacillus anthracis* and *Bacillus cereus* with repeated Cys-Xaa-Xaa motifs. *Biol. Direct.* 4:15. doi: 10.1186/1745-6150-4-15

Harms, J. M., Wilson, D. N., Schluenzen, F., Connell, S. R., Stachelhaus, T., Zaborowska, Z., et al. (2008). Translational regulation via L11: molecular switches on the ribosome turned on and off by Thiostrepton and Micrococcin. *Mol. Cell* 30, 26–38. doi: 10.1016/j.molcel.2008.01.009

Hayashi, S., Ozaki, T., Asamizu, S., Ikeda, H., Ômura, S., Oku, N., et al. (2014). Genome mining reveals a minimum gene set for the biosynthesis of 32-membered macrocyclic thiopeptides lactazoles. *Chem. Biol.* 21, 679–688. doi: 10.1016/j.chembiol.2014.03.008

Heddle, J. G., Blance, S. J., Zamble, D. B., Hollfelder, F., Miller, D. A., Wentzell, L. M., et al. (2001). The antibiotic microcin B17 is a DNA gyrase poison: characterisation of the mode of inhibition. *J. Mol. Biol.* 307, 1223–1234. doi: 10.1006/jmbi.2001.4562

Heffron, S. E., and Jurnak, F. (2000). Structure of an EF-Tu complex with a thiazolyl peptide antibiotic determined at 2.*35* a resolution: atomic basis for GE2270A inhibition of EF-Tu. *Biochemistry* 39, 37–45. doi: 10.1021/bi9913597

Hegde, S. S., Vetting, M. W., Roderick, S. L., Mitchenall, L. A., Maxwell, A., Takiff, H. E., et al. (2005). A fluoroquinolone resistance protein from *Mycobacterium tuberculosis* that mimics DNA. *Science* 308, 1480–1483. doi: 10.1126/science.1110699

Hudson, G. A., Burkhart, B. J., DiCaprio, A. J., Schwalen, C. J., Kille, B., Pogorelov, T. V., et al. (2019). Bioinformatic mapping of radical s-adenosylmethionine-dependent ribosomally synthesized and post-translationally modified peptides identifies new Cα, Cβ, and Cγ-linked thioether-containing peptides. *J. Am. Chem. Soc.* 141, 8228–8238. doi: 10.1021/jacs.9b01519

Hudson, G. A., Zhang, Z., Tietz, J. I., Mitchell, D. A., and van der Donk, W. A. (2015). In vitro biosynthesis of the core scaffold of the *Thiopeptide Thiomuracin*. *J. Am. Chem. Soc.* 137, 16012–16015. doi: 10.1021/jacs.5b10194

Huo, L., Rachid, S., Stadler, M., Wenzel, S. C., and Müller, R. (2012). Synthetic biotechnology to study and engineer ribosomal bottromycin biosynthesis. *Chem. Biol.* 19, 1278–1287. doi: 10.1016/j.chembiol.2012.08.013

Jarrad, A. M., Karoli, T., Blaskovich, M. A. T., Lyras, D., and Cooper, M. A. (2015). Clostridium difficile drug pipeline: challenges in discovery and development of new agents. *J. Med. Chem.* 58, 5164–5185. doi: 10.1021/jm5016846

Jeon, B. S., Wang, S. A., Ruszczycky, M. W., and Liu, H. W. (2017). Natural [4 + 2]-Cyclases. *Chem. Rev.* 117, 5367–5388. doi: 10.1021/acs.chemrev.6b00578

Just-Baringo, X., Albericio, F., and Álvarez, M. (2014). Thiopeptide antibiotics: retrospective and recent advances. *Mar. Drugs* 12, 317–351. doi: 10.3390/md12010317

Kautsar, S. A., Blin, K., Shaw, S., Navarro-Muñoz, J. C., Terlouw, B. R., van der Hooft, J. J. J., et al. (2019). MIBiG 2.0: a repository for biosynthetic gene clusters of known function. *Nucleic Acids Res.* 48, D454–D458. doi: 10.1093/nar/gkz882

Kazakov, T., Vondenhoff, G. H., Datsenko, K. A., Novikova, M., Metlitskaya, A., Wanner, B. L., et al. (2008). *Escherichia coli* peptidase A, B, or N can process translation inhibitor microcin C. *J. Bacteriol.* 190, 2607–2610. doi: 10.1128/JB.01956-07

Kulikovsky, A., Serebryakova, M., Bantysh, O., Metlitskaya, A., Borukhov, S., Severinov, K., et al. (2014). The molecular mechanism of aminopropylation of peptide-nucleotide antibiotic microcin C. *J. Am. Chem. Soc.* 136, 11168–11175. doi: 10.1021/ja505982c

Lee, J., Hao, Y., Blair, P. M., Melby, J. O., Agarwal, V., Burkhart, B. J., et al. (2013). Structural and functional insight into an unexpectedly selective N-methyltransferase involved in plantazolicin biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 110, 12954–12959. doi: 10.1073/pnas.1306101110

Li, J., Qu, X., He, X., Duan, L., Wu, G., Bi, D., et al. (2012). ThioFinder: a web-based tool for the identification of thiopeptide gene clusters in DNA sequences. *PLoS One* 7:e45878. doi: 10.1371/journal.pone.0045878

Li, J. W.-H., and Vederas, J. C. (2009). Drug discovery and natural products: end of an era or an endless frontier? *Science* 325, 161–165. doi: 10.1126/science.1168243

Li, Y. M., Milne, J. C., Madison, L. L., Kolter, R., and Walsh, C. T. (1996). From peptide precursors to oxazole and thiazole-containing peptide antibiotics: microcin B17 synthase. *Science* 274, 1188–1193. doi: 10.1126/science.274.5290.1188

Liu, N., Song, L., Liu, M., Shang, F., Anderson, Z., Fox, D. J., et al. (2016). Unique post-translational oxime formation in the biosynthesis of the azolemycin complex of novel ribosomal peptides from *Streptomyces* sp. FXJ1.264. *Chem. Sci.* 7, 482–488. doi: 10.1039/c5sc03021h

Mahanta, N., Liu, A., Dong, S., Nair, S. K., and Mitchell, D. A. (2018). Enzymatic reconstitution of ribosomal peptide backbone thioamidation. *Proc. Natl. Acad. Sci. U.S.A.* 2017, 22324. doi: 10.1073/pnas.1722324115

Mann, G., Huo, L., Adam, S., Nardone, B., Vendome, J., Westwood, N. J., et al. (2016). Structure and substrate recognition of the bottromycin maturation enzyme BotP. *Chembiochem* 17, 2286–2292. doi: 10.1002/cbic.201600406

Melby, J. O., Dunbar, K. L., Trinh, N. Q., and Mitchell, D. A. (2012). Selectivity, directionality, and promiscuity in peptide processing from a bacillus sp. *Al Hakam cyclodehydratase*. *J. Am. Chem. Soc.* 134, 5309–5316. doi: 10.1021/ja211675n

Metelev, M., Osterman, I. A., Ghilarov, D., Khabibullina, N. F., Yakimov, A., Shabalin, K., et al. (2017). Klebsazolicin inhibits 70S ribosome by obstructing the peptide exit tunnel. *Nat. Chem. Biol.* 13, 1129–1136. doi: 10.1038/nchembio.2462

Metelev, M., Serebryakova, M., Ghilarov, D., Zhao, Y., and Severinov, K. (2013). Structure of microcin B-like compounds produced by *Pseudomonas* syringae and species specificity of their antibacterial action. *J. Bacteriol.* 195, 4129–4137. doi: 10.1128/JB.00665-13

Metlitskaya, A., Kazakov, T., Kommer, A., Pavlova, O., Praetorius-Ibba, M., Ibba, M., et al. (2006). Aspartyl-tRNA synthetase is the target of peptide nucleotide antibiotic Microcin C. *J. Biol. Chem.* 281, 18033–18042. doi: 10.1074/jbc.M513174200

Mirdita, M., Steinegger, M., and Söding, J. (2019). MMseqs2 desktop and local web server app for fast, interactive sequence searches. *Bioinformatics* 35, 2856–2858. doi: 10.1093/bioinformatics/bty1057

Molloy, E. M., Cotter, P. D., Hill, C., Mitchell, D. A., and Ross, R. P. (2011). Streptolysin S-like virulence factors: the continuing sagA. *Nat. Rev. Microbiol.* 9, 670–681. doi: 10.1038/nrmicro2624

Moloney, M. G. (2016). Natural products as a source for novel antibiotics. *Trends Pharmacol. Sci.* 37, 689–701. doi: 10.1016/j.tips.2016.05.001

Navarro-Muñoz, J. C., Selem-Mojica, N., Mullowney, M. W., Kautsar, S. A., Tryon, J. H., Parkinson, E. I., et al. (2020). A computational framework to explore large-scale biosynthetic diversity. *Nat. Chem. Biol.* 16, 60–68. doi: 10.1038/s41589-019-0400-9

Nissen, P., Thirup, S., Kjeldgaard, M., and Nyborg, J. (1999). The crystal structure of Cys-tRNACys-EF-Tu-GDPNP reveals general and specific features in the ternary complex and in tRNA. *Structure* 7, 143–156. doi: 10.1016/s0969-2126(99)80021-5

Novikova, M., Metlitskaya, A., Datsenko, K., Kazakov, T., Kazakov, A., Wanner, B., et al. (2007). The *Escherichia coli* Yej transporter is required for the uptake of translation inhibitor microcin C. *J. Bacteriol.* 189, 8361–8365. doi: 10.1128/JB.01028-07

O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., et al. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* 44, D733–D745. doi: 10.1093/nar/gkv1189

Otaka, T., and Kaji, A. (1976). Mode of action of bottromycin A2. Release of aminoacyl- or peptidyl-tRNA from ribosomes. *J. Biol. Chem.* 251, 2299–2306.

Otaka, T., and Kaji, A. (1981). Mode of action of bottromycin A2: effect on peptide bond formation. *FEBS Lett.* 123, 173–176. doi: 10.1016/0014-5793(81)80280-3

Otaka, T., and Kaji, A. (1983). Mode of action of bottromycin A2: effect of bottromycin A2 on polysomes. *FEBS Lett.* 153, 53–59. doi: 10.1016/0014-5793(83)80118-5

Ozaki, T., Kurokawa, Y., Hayashi, S., Oku, N., Asamizu, S., Igarashi, Y., et al. (2016). Insights into the Biosynthesis of Dehydroalanines in Goadsporin. *Chembiochem* 17, 218–223. doi: 10.1002/cbic.201500541

Panter, F., Krug, D., Baumann, S., and Müller, R. (2018). Self-resistance guided genome mining uncovers new topoisomerase inhibitors from myxobacteria. *Chem. Sci.* 9, 4898–4908. doi: 10.1039/c8sc01325j

Parmeggiani, A., Krab, I. M., Okamura, S., Nielsen, R. C., Nyborg, J., and Nissen, P. (2006). Structural basis of the action of pulvomycin and GE2270 A on elongation factor Tu. *Biochemistry* 45, 6846–6857. doi: 10.1021/bi0525122

Polikanov, Y. S., Aleksashin, N. A., Beckert, B., and Wilson, D. N. (2018). The mechanisms of action of ribosome-targeting peptide antibiotics. *Front. Mol. Biosci.* 5:48. doi: 10.3389/fmolb.2018.00048

Repka, L. M., Chekan, J. R., Nair, S. K., and van der Donk, W. (2017). Mechanistic understanding of lanthipeptide biosynthetic enzymes. *Chem. Rev.* 117, 5457–5520. doi: 10.1021/acs.chemrev.6b00591

Rodnina, M. V., Savelsbergh, A., Matassova, N. B., Katunin, V. I., Semenkov, Y. P., and Wintermeyer, W. (1999). Thiostrepton inhibits the turnover but not the GTPase of elongation factor G on the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* 96, 9586–9590. doi: 10.1073/pnas.96.17.9586

Roush, R. F., Nolan, E. M., Löhr, F., and Walsh, C. T. (2008). Maturation of an *Escherichia coli* ribosomal peptide antibiotic by ATP-consuming N-P bond formation in microcin C7. *J. Am. Chem. Soc.* 130, 3603–3609. doi: 10.1021/ja7101949

Santos-Aberturas, J., Chandra, G., Frattaruolo, L., Lacret, R., Pham, T. H., Vior, N. M., et al. (2019). Uncovering the unexplored diversity of thioamidated ribosomal peptides in Actinobacteria using the RiPPER genome mining tool. *Nucleic Acids Res.* 47, 4624–4637. doi: 10.1093/nar/gkz192

Schramma, K. R., Bushin, L. B., and Seyedsayamdost, M. R. (2015). Structure and biosynthesis of a macrocyclic peptide containing an unprecedented lysine-to-tryptophan crosslink. *Nat. Chem.* 7, 431–437. doi: 10.1038/nchem.2237

Schwalen, C. J., Hudson, G. A., Kille, B., and Mitchell, D. A. (2018). Bioinformatic expansion and discovery of thiopeptide antibiotics. *J. Am. Chem. Soc.* 140, 9494–9501. doi: 10.1021/jacs.8b03896

Schwalen, C. J., Hudson, G. A., Kosol, S., Mahanta, N., Challis, G. L., and Mitchell, D. A. (2017). In vitro biosynthetic studies of bottromycin expand the enzymatic capabilities of the YcaO superfamily. *J. Am. Chem. Soc.* 139, 18154–18157. doi: 10.1021/jacs.7b09899

Selva, E., Beretta, G., Montanini, N., Saddler, G. S., Gastaldo, L., Ferrari, P., et al. (1991). Antibiotic GE2270 a: a novel inhibitor of bacterial protein synthesis. I.

isolation and characterization. *J. Antibiot.* 44, 693–701. doi: 10.7164/antibiotics.44.693

Serebryakova, M., Tsibulskaya, D., Mokina, O., Kulikovsky, A., Nautiyal, M., Van Aerschot, A., et al. (2016). A trojan-horse peptide-carboxymethyl-cytidine antibiotic from *Bacillus amyloliquefaciens*. *J. Am. Chem. Soc.* 138, 15690–15698. doi: 10.1021/jacs.6b09853

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303

Shim, Y. Y., Young, L. W., Arnison, P. G., Gilding, E., and Reaney, M. J. T. (2015). Proposed systematic nomenclature for orbitides. *J. Nat. Prod.* 78, 645–652. doi: 10.1021/np500802p

Shimamura, H., Gouda, H., Nagai, K., Hirose, T., Ichioka, M., Furuya, Y., et al. (2009). Structure determination and total synthesis of bottromycin A2: a potent antibiotic against MRSA and VRE. *Angew. Chemie - Int. Ed.* 48, 914–917. doi: 10.1002/anie.200804138

Sikandar, A., Franz, L., Melse, O., Antes, I., and Koehnke, J. (2019). Thiazoline-specific amidohydrolase PurAH is the gatekeeper of bottromycin biosynthesis. *J. Am. Chem. Soc.* 141, 9748–9752. doi: 10.1021/jacs.8b12231

Söding, J., Biegert, A., and Lupas, A. N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* 33, W244–W248. doi: 10.1093/nar/gki408

Süssmuth, R. D., and Mainz, A. (2017). Nonribosomal peptide synthesis—principles and prospects. *Angew. Chem. Int. Ed.* 56, 3770–3821. doi: 10.1002/anie.201609079

Tanaka, N., Sashikata, K., Yamaguchi, H., and Umezawa, H. (1966). Inhibition of protein synthesis by bottromycin A2 and its hydrazide. *J. Biochem.* 60, 405–410. doi: 10.1093/oxfordjournals.jbchem.a128451

Tietz, J. I. T., Schwalen, C. J., Patel, P. S., Maxon, T., Blair, P. M., Tai, H.-C., et al. (2017). A new genome mining tool redefines the lasso peptide biosynthetic landscape. *Nat. Chem. Biol.* 13, 470–478. doi: 10.1038/nchembio.2319

Travin, D. Y., Metelev, M., Serebryakova, M., Komarova, E. S., Osterman, I. A., Ghilarov, D., et al. (2018). Biosynthesis of translation inhibitor klebsazolicin proceeds through heterocyclization and n-terminal amidine formation catalyzed by a single YcaO enzyme. *J. Am. Chem. Soc.* 140, 5625–5633. doi: 10.1021/jacs.8b02277

Travin, D. Y., Watson, Z. L., Metelev, M., Ward, F. R., Osterman, I. A., Khven, I. M., et al. (2019). Structure of ribosome-bound azole-modified peptide phazolicin rationalizes its species-specific mode of bacterial translation inhibition. *Nat. Commun.* 10:4563. doi: 10.1038/s41467-019-12589-5

Tsibulskaya, D., Mokina, O., Kulikovsky, A., Piskunova, J., Severinov, K., Serebryakova, M., et al. (2017). The product of *Yersinia pseudotuberculosis* mcc operon is a peptide-cytidine antibiotic activated inside producing cells by the tldd/e protease. *J. Am. Chem. Soc.* 139, 16178–16187. doi: 10.1021/jacs.7b07118

Vetting, M. W., Hegde, S. S., Wang, M., Jacoby, G., Hooper, D. C., and Blanchard, J. S. (2011). Structure of QnrB1, a plasmid-mediated fluoroquinolone resistance factor. *J. Biol. Chem.* 286, 25265–25273. doi: 10.1074/jbc.M111.226936

Waisvisz, J. M., van der Hoeven, M. G., van Peppen, J., and Zwennis, W. C. M. (1957). Bottromycin. I. a new sulfur-containing antibiotic. *J. Am. Chem. Soc.* 79, 4520–4521. doi: 10.1021/ja01573a072

Wilson, D. N. (2014). Ribosome-targeting antibiotics and mechanisms of bacterial resistance. *Nat. Rev. Microbiol.* 12, 35–48. doi: 10.1038/nrmicro3155

Wright, G. D. (2017). Opportunities for natural products in 21st century antibiotic discovery. *Nat. Prod. Rep.* 34, 694–701. doi: 10.1039/c7np00019g

Young, T. S., Dorrestein, P. C., and Walsh, C. T. (2012). Codon randomization for rapid exploration of chemical space in thiopeptide antibiotic variants. *Chem. Biol.* 19, 1600–1610. doi: 10.1016/j.chembiol.2012.10.013

Zheng, Q., Fang, H., and Liu, W. (2017). Post-translational modifications involved in the biosynthesis of thiopeptide antibiotics. *Org. Biomol. Chem.* 15, 3376–3390. doi: 10.1039/c7ob00466d

Ziemert, N., Alanjary, M., and Weber, T. (2016). The evolution of genome mining in microbes - a review. *Nat. Prod. Rep.* 33, 988–1005. doi: 10.1039/c6np00025h

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

REVIEW
published: 17 April 2020
doi: 10.3389/fgene.2020.00262

The Variety in the Common Theme of Translation Inhibition by Type II Toxin–Antitoxin Systems

*Dukas Jurėnas[1] and Laurence Van Melderen[2]**

[1] Laboratoire d'Ingénierie des Systèmes Macromoléculaires, Institut de Microbiologie de la Méditerranée, CNRS, Aix-Marseille Université, Marseille, France, [2] Cellular and Molecular Microbiology, Faculté des Sciences, Université libre de Bruxelles, Gosselies, Belgium

Type II Toxin–antitoxin (TA) modules are bacterial operons that encode a toxic protein and its antidote, which form a self-regulating genetic system. Antitoxins put a halter on toxins in many ways that distinguish different types of TA modules. In type II TA modules, toxin and antitoxin are proteins that form a complex which physically sequesters the toxin, thereby preventing its toxic activity. Type II toxins inhibit various cellular processes, however, the translation process appears to be their favorite target and nearly every step of this complex process is inhibited by type II toxins. The structural features, enzymatic activities and target specificities of the different toxin families are discussed. Finally, this review emphasizes that the structural folds presented by these toxins are not restricted to type II TA toxins or to one particular cellular target, and discusses why so many of them evolved to target translation as well as the recent developments regarding the role(s) of these systems in bacterial physiology and evolution.

Keywords: toxins, translation, persistence, programmed cell death, mobile genetic elements

OPEN ACCESS

Edited by:
Michael Ibba,
The Ohio State University, United States

Reviewed by:
Christine M. Dunham,
Emory University, United States
Dominique Belin,
Université de Genève, Switzerland

****Correspondence:***
Laurence Van Melderen
lvmelder@ulb.ac.be;
laurence.van.melderen@ulb.ac.be

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *15 October 2019*
Accepted: *05 March 2020*
Published: *17 April 2020*

Citation:
Jurėnas D and Van Melderen L (2020) The Variety in the Common Theme of Translation Inhibition by Type II Toxin–Antitoxin Systems. Front. Genet. 11:262. doi: 10.3389/fgene.2020.00262

INTRODUCTION

Bacterial toxin-antitoxin (TA) systems are generally composed of a toxic protein and its inhibitor. These small modules were originally discovered on plasmids in the 1980s where they were found to promote plasmid maintenance in growing bacterial populations (Karoui et al., 1983; Ogura and Hiraga, 1983; Jaffe et al., 1985; Gerdes et al., 1986; Hiraga et al., 1986). TA modules are classified into different types depending on the nature and mode of action of the antitoxins, as the toxins are always proteins. Antitoxins are either small RNAs that block translation of the toxin mRNA (type I) (Gerdes and Wagner, 2007) or sequesters the toxic protein (type III) (Fineran et al., 2009; Blower et al., 2011) or proteins that inhibit the activity of the toxin through direct interaction (type II) (Tam and Kline, 1989; Kamada et al., 2003; Takagi et al., 2005) or antagonize the toxic activity on the target, without any direct interaction with the toxins (type IV) (Brown and Shaw, 2003). This review will focus on type II systems. These elements are not only found in plasmids but also in other types of mobile genetic elements (such as phages and ICEs) as well as in chromosomes (see e.g., Anantharaman and Aravind, 2003; Pandey and Gerdes, 2005; Guglielmini and Van Melderen, 2011; Leplae et al., 2011; Ramisetty et al., 2016; Coray et al., 2017). While the roles of TAs, when located in mobile genetic elements, are reminiscent to that on plasmids, i.e., maintenance (Szekeres et al., 2007; Wozniak and Waldor, 2009; Huguet et al., 2016), the roles of chromosomally-encoded systems remains a largely debated topic in the field. These systems have been involved in the

adaptation to adverse conditions and are considered to be stress response modules (Hayes and Van Melderen, 2011; Page and Peti, 2016; Harms et al., 2018), with a mainstream model proposing that TA systems are essential effectors of persistence to antibiotics (Gerdes and Maisonneuve, 2012). However, seminal papers supporting this hypothesis are now being retracted (Maisonneuve et al., 2018a,b; Germain et al., 2019) and contradictory data are being published (Harms et al., 2017; Shan et al., 2017; Goormaghtigh et al., 2018a; Pontes and Groisman, 2019). The involvement of TA systems in the persistence phenomenon was based on the observation that successive deletions of 10 TA systems in *Escherichia coli* lead to a gradual decrease of persistence frequency in the presence of lethal doses of ampicillin or ciprofloxacin (Maisonneuve et al., 2011). In subsequent studies, time-lapse microscopy experiments with *E. coli* strains containing fluorescent reporters revealed that cells that are able to recover from an ampicillin treatment (namely persister cells) are those in which the ppGpp level is high and TA systems are activated (Maisonneuve et al., 2013). Furthermore, the same group proposed that the HipBA system is the major regulator of persistence as the HipA toxin, by phosphorylating glutamyl-tRNA synthetase (see below), will trigger ppGpp production and the activation of the 10 other TA systems (Germain et al., 2015). We, along with other researchers, have identified major problems both with the *E. coli* strains and fluorescent reporters that were used in these studies (Harms et al., 2017; Goormaghtigh et al., 2018a). First, the strain deleted for the 10 TA systems is lysogenized with several copies of the Phi80 phages (Harms et al., 2017; Goormaghtigh et al., 2018a). This explains why that strain presents a lower persistence frequency to ciprofloxacin. Indeed, activation of the SOS response by fluoroquinolones will lead to lambdoid prophage activation and a strong decrease in viability. Second, the fluorescent reporters used to monitor ppGpp levels and TA activation are likely to not be functional, either forming protein aggregates or not being more fluorescent than the fluorescence background of control strains without the reporter (Goormaghtigh et al., 2018a). In an effort to solve the issue of TAs and persistence, we, along with other researchers, constructed a strain in which the 10 TA systems are deleted, devoid of any phage contaminants, and showed that this strain presents the same level of persistence to ampicillin or ofloxacin as the wild-type strain (Harms et al., 2017; Goormaghtigh et al., 2018a). Moreover, newly designed reporters monitoring the activation of the *yefM-yoeB* TA system allowed us to show that there is no correlation between persister cells and the activation of this TA system (Goormaghtigh et al., 2018a). Moreover, we recently showed that another TA system, MqsrA (see below), that was thought to be a global regulator involved in stress responses and biofilm formation, does not appear to play any significant role neither in oxidative or bile stresses nor in macrocolony formation (Fraikin et al., 2019). Therefore, our data strongly argue against the idea that TA systems are pivotal elements of antibiotic persistence. This basically leaves the principal questions in the field open (for a recent review see Fraikin et al., 2020). Conditions in which TA systems are activated and what the outcomes of such activations are, still remain undetermined.

Type II toxins are very diverse in their molecular mode of action, however, almost all the families described to date comprise toxins that target protein synthesis (Harms et al., 2018). Several major mechanisms of translation inhibition by type II toxins can be distinguished: RNA hydrolysis of (i) solvent exposed RNAs, (ii) mRNAs in ribosomes, (iii) rRNAs, (iv) tRNAs or interference with the tRNA machinery, where in addition to the above mentioned hydrolysis of tRNAs, toxicity is exhibited by the modification of tRNA cargo (v) and the inactivation of enzymes that service the tRNAs (vi) i.e., phosphorylation of aminoacyl-tRNA synthetases that charge the tRNAs or EF-Tu which delivers the tRNAs to the ribosome (**Figure 1**). Most type II toxin activities lead to the general inhibition of protein synthesis and the subsequent inhibition of growth. In this review, we follow this classification to discuss the specificity and evolution of the different families of toxins from TA systems. We discuss the consequences of toxin-mediated translation inhibition on cell physiology and phenotypes and raise questions about their biological functions in light of recent discoveries.

HYDROLYSIS OF RNAs

MazF Toxins

MazF toxins are RNA endonucleases that exhibit cleavage specificity to sequences spanning from three to seven bases (**Figure 2**). Most of the MazF toxins prefer U upstream of cleavage (at position −1) and AC downstream of cleavage (at positions +1 and +2, respectively) with less stringency outside of the two to four main recognized bases (**Figure 2**). The *E. coli* MazF cleaves in the coding as well as untranslated regions of mRNAs, independent from the reading frame, and in rRNA precursors (Mets et al., 2017; Culviner and Laub, 2018; Mets et al., 2019). Slight changes in sequence specificity of different MazF enzymes from different bacteria, correspond well to their amino acid sequence similarity (**Figure 2**). A particular member of the MazF family, the MazF-mt9 toxin from *Mycobacterium tuberculosis*, cleaves a tRNA substrate (Schifano et al., 2016), similar to the VapC toxins that will be discussed later in this review.

The structures of many MazF toxins alone or in complex with an mRNA substrate or with their cognate antitoxins have been solved (Hargreaves et al., 2002; Kamada et al., 2003; Simanshu et al., 2013; Zorzini et al., 2014, 2016; Ahn et al., 2017; Chen et al., 2017; Hoffer et al., 2017). The MazF monomer consists of 2 beta-sheets composed of antiparallel beta-strands linked by three or four small alpha-helices (**Figure 3**). The residues constituting the active site are located on the β1-β2 and β3-β4 linkers (Kamada et al., 2003; Simanshu et al., 2013; Zorzini et al., 2016). Although MazF toxins possess what is known as the SH3-barrel-fold, they are not related to other members of this fold (Anantharaman and Aravind, 2003). Two MazF subunits form a dimer with an extensive dimeric interface. A concave positively charged groove at the interface between the two subunits of the MazF dimer binds RNA in an extended alignment with bases facing upwards toward the groove (Simanshu et al., 2013; Zorzini et al., 2016). Structural studies have shown that

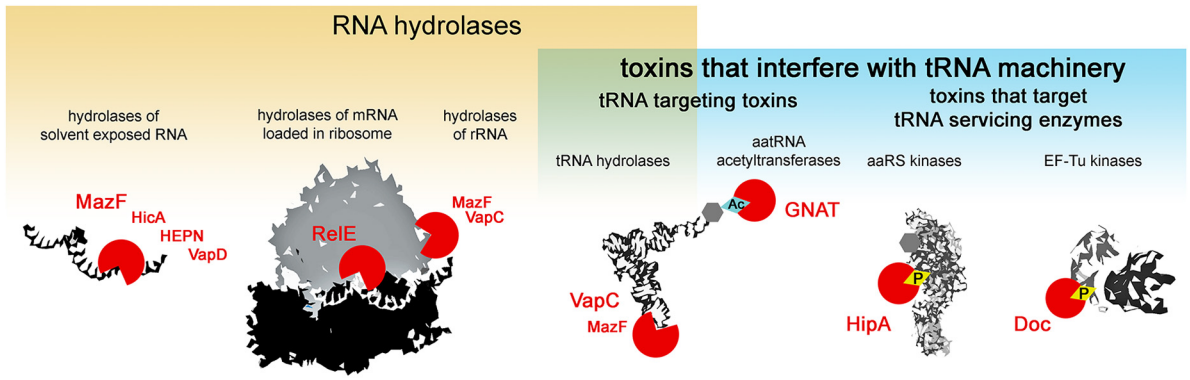

FIGURE 1 | Activities of type II TA toxins. Cellular targets of the toxins are depicted in black and white; toxins are depicted as red circles – open circles for toxins hydrolyzing chemical bonds, circles with diamond for toxins transferring chemical groups on the targets. Ac stands for acetylation, and P for phosphorylation.

a MazF dimer binds to one mRNA molecule and covers at least seven bases (Simanshu et al., 2013; Zorzini et al., 2016), explaining the extent of possible recognition and necessity of an unstructured RNA substrate, i.e., solvent exposed bases. MazF cleavage leaves 2′, 3′-cyclic phosphate at the 3′-end and 5′-OH group at the 5′-end of the cleavage site, which are used as a signature to study the cleavage products generated by the *in vivo* overexpression of MazF toxins (Schifano et al., 2014; Mets et al., 2017).

In addition to cleaving mRNAs, the *E. coli* MazF toxin was shown to cleave the 16S rRNA (Vesper et al., 2011). It was proposed that rRNA-cleavage generates specialized ribosomes able to translate specific pools of mRNAs constituting a regulon that is required to cope with various stresses (Vesper et al., 2011; Sauert et al., 2016). The model relies on data showing that *E. coli* MazF cleaves the 16S rRNA in the decoding center before ACA sequences at nucleotide positions 1396 and 1500 ($\downarrow^{1396}$ACA and $\downarrow^{1500}$ACA) within the 30S ribosomal subunit. It was speculated that MazF-mediated cleavage removes the anti-SD sequence, thereby generating specialized ribosomes that are able to translate specialized MazF-processed leaderless mRNAs that also lack SD (Vesper et al., 2011; Sauert et al., 2016). However, these particular 16S rRNA bases are paired in the 30S subunits, making the MazF-dependent cleavage very unlikely. It was later shown that indeed, MazF cleaves the rRNAs in their precursor states, before the 30S subunit biogenesis, and at multiple sites (Mets et al., 2017; Culviner and Laub, 2018). This is in agreement with structural data showing that MazF binds to single-stranded RNA (ssRNA) and interacts with the bases (Simanshu et al., 2013; Zorzini et al., 2016). In fact, fragmentation of rRNAs as well as mRNAs including those coding for ribosomal proteins leads to the accumulation of aberrant ribosomal subunits and generates irregular particles with fragmented rRNA upon MazF expression (Mets et al., 2017). Accordingly, the synthesis of the vast majority of cellular proteins drop in response to MazF cleavage, without enrichment of any specific functional protein groups. Proteomic studies upon MazF expression did not find any 'death' or 'survival' proteins (Mets et al., 2019), associated to the proposed programmed cell death pathway (Amitai et al., 2009).

The TA-mediated programmed cell death (PCD) theory proposed that some TA systems, in particular MazEF, serve as built-in suicide modules (Engelberg-Kulka et al., 2004). Multiple different and unrelated stressing conditions, such as amino acids starvation and elevated ppGpp (guanosine tetraphosphate) concentrations, antibiotic treatments, high temperature, H_2O_2 treatment or phage infections were proposed to trigger MazF-dependent PCD. All these conditions would lead to the inhibition of *mazEF* expression. Since the MazE antitoxin is unstable and degraded by ATP-dependent proteases, this would liberate MazF and provoke the death of a large subpopulation of cells (Engelberg-Kulka et al., 2004). Within the frame of this pathway, it has been suggested that MazF itself is induced by a 'quorum sensing' peptide called extracellular death factor (EDF). This pentapeptide was identified to have the NNWNN sequence (Kolodkin-Gal et al., 2007) and is thought to amplify the endoribonucleolytic activity of MazF and other homologs and prevent the interaction of MazF toxins with their cognate antitoxins (Belitsky et al., 2011). However, alongside the PCD theory, the involvement of the EDF peptide in MazF-mediated PCD is questionable. First, in the original paper that describes the activity of the EDF peptide, the authors observed a drastic decrease in bacterial viability upon the addition of 2.5 ng/ml (4 nM) of EDF to growing *E. coli* cultures. This effect was also observed in MazF-deleted strains although at higher concentrations, starting from 200 ng/ml (300 nM) (Kolodkin-Gal et al., 2007). In subsequent publications, the authors showed that *in vitro* MazF ribonucleolytic activity was enhanced by only 26% by the addition of 1.5 μM of EDF (Belitsky et al., 2011). Further addition of up to 7.5 μM of EDF increased MazF activity to a maximum of 57% (Belitsky et al., 2011). Such a high EDF concentration, corresponding to a 300-fold molar excess, was used to show the disruption of the MazE-MazF complex *in vitro*. Assuming that EDF is produced *in vivo* under physiological conditions, it is very unlikely that such high concentrations will be reached. Moreover, *in vivo*, viability of bacteria is highly affected at EDF concentrations that are by several logs lower (4 nM versus 7.5 μM). The same group later discovered more EDF-like peptides (6 aa or longer) in

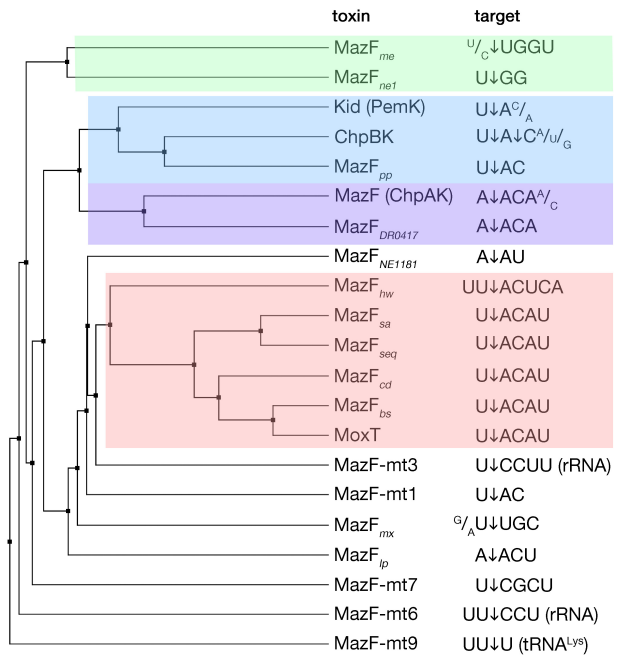

FIGURE 2 | MazF toxins and their specificity of RNA cleavage. MazF protein sequences were aligned and an average distance tree was build (BLOSUM 62, JalView) (Waterhouse et al., 2009). Substrate specificity is indicated (right). The cleavage position of the substrate is indicated by an arrow. Subsets of toxins sharing similar target sequences are boxed with different colors. Protein identifiers and cleavage specificity were taken from: *E. coli* MazF (ChpAK) (NP_417262.1) (Culviner and Laub, 2018; Mets et al., 2019), ChpBK (NP_418646.1) (Zhang et al., 2005), Kid (PemK) (YP_003937673.1) (Munoz-Gomez et al., 2004; Zhang et al., 2004); *Bacillus subtilis* MazF$_{bs}$ (AOR96854.1) (Park et al., 2011); *Bacillus anthracis* MoxT (NP_842807.1) (Verma and Bhatnagar, 2014); *Staphylococcus aureus* MazF$_{sa}$ (BBJ19047.1) (Zhu et al., 2009); *Staphylococcus equorum* MazF$_{seq}$ (AFV93478.1) (Schuster et al., 2013); *Clostridium difficile* MazF$_{cd}$ (YP_001089981.1) (Rothenbacher et al., 2012); *Legionella pneumophila* MazF$_{lp}$ (CCD10720.1) (Shaku et al., 2018); *Deinococcus radiodurans* MazF$_{DR0417}$ (AAF09995.1) (Miyamoto et al., 2017); *Haloquadratum walsbyi* MazF$_{hw}$ (WP_048066888.1) (Yamaguchi et al., 2012); *Pseudomonas putida* MazF$_{pp}$ (NP_742932.1) (Miyamoto et al., 2016a); *Myxococcus xanthus* MazF$_{mx}$ (SDX28280.1) (Nariya and Inouye, 2008); *Methanohalobium evestigatum* MazF$_{me}$ (WP_013195679.1); *Nitrosomonas europaea* MazF$_{ne1}$ (WP_011111532.1) (Miyamoto et al., 2018), MazF$_{NE1181}$ (CAD85092.1) (Miyamoto et al., 2016b); *Mycobacterium tuberculosis* MazF-mt1 (NP_217317.1) (Zhu et al., 2006), MazF-mt3 (NP_216507.1) (Zhu et al., 2008; Schifano et al., 2014), MazF-mt6 (NP_215618.1) (Schifano et al., 2013), MazF-mt7 (NP_216011.1) (Zhu et al., 2008), MazF-mt9 (YP_004837055.2) (Barth et al., 2019).

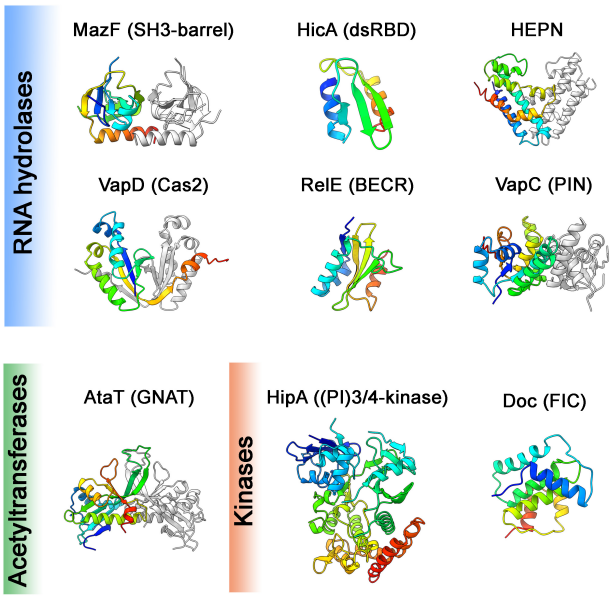

FIGURE 3 | Toxin structures. Structures of the toxins from different families representing different structural folds (noted in brackets) are shown. These domains are common in proteins of the defense and offense systems (HEPN, Cas2, BECR, FIC), RNA processing (dsRBD, PIN), eukaryotic cell signaling (SH3-barrel, (PI)3/4-kinase) or various other functions (GNAT). Structures are colored in rainbow from N-terminus (blue) to C-terminus (red), in cases of dimers the second monomer is in gray. Structures were visualized using ChimeraX, PDB codes were following: *E. coli* MazF:3nfc; *E. coli* HicAB:6hpb (Manav et al., 2019); *S. oneidensis* SO_3166-SO3165 (HEPN): 5yep (Jia et al., 2018); *H. pylori* HP0315 (VapD): 3ui3 (Kwon et al., 2012); *S. flexneri* VapC^{D7A}: 5ecw (Xu et al., 2016); *E. coli* RelE^{R81A,R83A}: 2kc9 (Li et al., 2009); *E. coli* HipA^{S150A}: 3tpb (Schumacher et al., 2012); *E. coli* AtaT^{Y144F}: 6gtp (Jurenas et al., 2019); prophage P1 Phd-Doc: 3k33 (Garcia-Pino et al., 2010). In the cases of structures in complexes with antitoxin, the coordinates of antitoxin were deleted.

Bacillus subtilis and *Pseudomonas aeruginosa*, however all of these peptides were only active in combination with rifampicin (Kolodkin-Gal et al., 2007; Kumar et al., 2013). Although in the original studies, rifampicin was used to induce MazF-mediated PCD (Kolodkin-Gal et al., 2007; Kumar et al., 2013), other studies did not find any MazF-related effect of rifampicin, nor any MazF-mediated PCD (Tsilibaris et al., 2007; Ramisetty et al., 2016). Thus, the cellular functions of MazF remain an open question.

Probably, the most intriguing case is that of *M. tuberculosis* which encodes 10 MazF toxins among its large TA arsenal (Sala et al., 2014). Out of those, the MazF-mt3 and MazF-mt6 toxins cleave mainly mRNAs at the U↓CCUU and UU↓CCU sequences, respectively. In addition, MazF-mt3 cleaves the anti-SD sequence of the 16S rRNA and both the MazF-mt3 and MazF-mt6 cleave the 23S rRNA loop 70 (L70) opening to the ribosomal A site (Schifano et al., 2013, 2014; Hoffer et al., 2017). It has been shown that MazF-mt6 can cleave the 23S rRNA in mature 50S subunit, albeit with 30% efficiency as compared to free RNA fragment coding for 23S L70 sequence (Hoffer et al., 2017). This indicates that like other MazF toxins, MazF-mt6 cleaves the solvent exposed target sequence. Alternatively, cleavage might be possible in rRNA precursors as described for *E. coli* MazF (Mets et al., 2017; Culviner and Laub, 2018). Another MazF toxin from *M. tuberculosis*, MazF-mt9, cleaves the $tRNA^{Lys43-UUU}$ in its anticodon sequence (Schifano et al., 2016; Barth et al., 2019). Since $tRNA^{Lys19-CUU}$ does not compensate for $tRNA^{Lys43-UUU}$, ribosomes stall at the AAA codon in MazF-mt9 overexpression conditions (Barth et al., 2019). Ribosome stalled transcripts are further cleaved by specific RNAses such as RNAse J and are therefore eliminated (Barth et al., 2019). Since the AAA codon is rare in the GC-rich *M. tuberculosis* (5.3/1000), the authors speculate that expression of MazF-mt9

generates a specific proteome consisting of the proteins whose genes are poor in AAA Lys codons (Barth et al., 2019). However, it remains unclear whether genes devoid of the AAA codon would be sufficient to generate a functional proteome.

Interestingly, MazF toxins share high structural similarity to the F plasmid CcdB toxin (Loris et al., 1999; Hargreaves et al., 2002; Gogos et al., 2003; Kamada et al., 2003) that binds to the GyrA subunits of DNA-gyrase and induces double-strand breaks and the SOS response (Bernard and Couturier, 1992; Bernard et al., 1993; Dao-Thi et al., 2005). Despite their structural similarity, the two toxins bind their substrates using different sites on the toxin dimer interface – CcdB binds GyrA *via* the α4 helices, while, MazF recognizes RNA using the β1-β2, β3-β4, β4-β5 loops and a short α1 helix (Dao-Thi et al., 2005; Zorzini et al., 2016). On the other hand, the activity of MazF and CcdB toxins is regulated by their cognate antitoxins in a common way, which further supports the hypothesis of a common ancestor (Zorzini et al., 2016).

HicA Toxins

HicA toxin and its cognate HicB antitoxin owe their gene names to a genetic locus linked to the pilus gene cluster (hif contiguous) in *Haemophilus influenzae* (Mhlanga-Mutangadura et al., 1998). HicA toxins possess small ∼50 amino acid double-stranded RNA binding domains (dsRBD). HicA folds into a three-stranded antiparallel beta-sheet flanked by two alpha helices that reside on one side of the sheet (**Figure 3**) (Makarova et al., 2006; Butt et al., 2014). The positively charged surface is predicted to bind RNA and the catalytic histidine residue located in the β2 strand is required for RNase activity (Bibi-Triki et al., 2014; Butt et al., 2014; Kim et al., 2018). The *E. coli* HicA toxin cleaves mRNAs and tmRNA *in vivo* independently of translation and no consensus of cleavage was reported (Jorgensen et al., 2009). *Yersinia pestis* HicA3 was shown to degrade *in vitro* transcribed mRNAs (Bibi-Triki et al., 2014) and *Sinorhizobium meliloti* HicA degrades purified rRNA (Thomet et al., 2019). However, there is not enough data available to be able to conclude what the precise targets of HicA toxins are *in vivo*.

Outside of the TA context, it has been shown that the dsRBD domain binds double-stranded RNA where the α1 helix interacts with the minor groove and the α2 helix with the major groove of RNA molecules (Ryter and Schultz, 1998). All the HicA toxins studied to date hydrolyze RNA in addition to binding (Jorgensen et al., 2009; Bibi-Triki et al., 2014; Thomet et al., 2019), however there are no details on the molecular substrate binding and hydrolysis.

The typical antitoxin partner HicB protein comprises DNA-binding domain fused to a degenerated RNAse H-fold (Makarova et al., 2006). These two domains (dsRBD and RNAse H) are also found in the architecture of eukaryotic RNA interference (RNAi) machinery (Makarova et al., 2006).

HEPN-Fold Toxins

The HEPN (Higher Eukaryotes and Prokaryotes Nucleotide-binding domain) superfamily contains proteins that have all-alpha helical catalytic domains. Outside of the TA context, HEPN-domain proteins typically have metal-independent endonuclease activities, although some only bind RNA without degrading it (Anantharaman et al., 2013). HEPN domains are shared between TA and prokaryotic defense systems, such as abortive infection modules, restriction-modification systems and CRISPR-Cas systems, as well as eukaryotic antiviral, antitransposon systems and rRNA processing enzymes (Anantharaman et al., 2013). In the TA context, HEPN is frequently found in association to MNT (minimal nucleotidyltransferase) domain proteins. *Shewanella oneidensis* SO_3166 toxin possesses a HEPN domain and its cognate antitoxin SO_3165 an MNT domain (Yao et al., 2015). The SO_3166 toxin cleaves mRNA, but not rRNA or tRNA *in vitro*, however, the sequence specificity has not been determined. The toxin has a typical all-alpha helical HEPN fold and conserved Arg-(4-6X)-His motif, and forms a dimer with a potential composite active site in a central cleft that could accommodate RNA substrate (**Figure 3**) (Jia et al., 2018).

RnlA (or RNase LS)-like toxins that exhibit RNAse activity (Otsuka et al., 2007) also contain a catalytic domain belonging to the HEPN superfamily (Anantharaman et al., 2013). RNA cleavage by RnlA was shown to be dependent on translation (Otsuka and Yonesaki, 2012), more specifically on translation termination (Yamanishi and Yonesaki, 2005). RnlA induces sequence non-specific mRNA cleavage more frequently occurring 3′ to pyrimidines (Kai and Yonesaki, 2002). Homologous toxin LsoA shares the same cleavage pattern, however, cleavage sites are not identical (Otsuka and Yonesaki, 2012). The RnlA and LsoA toxins were shown to be part of functional TA systems with their cognate antitoxins RnlB and LsoB, respectively (Koga et al., 2011; Otsuka and Yonesaki, 2012). Interestingly, it was shown that the RnlA and LsoA toxins have a common antitoxin, Dmd, encoded by the T4 phage (Otsuka and Yonesaki, 2012). This antitoxin does not share any sequence similarity with the homologous but not interchangeable RnlB and LsoB antitoxins. T4 phages devoid of the Dmd-encoding gene are unable to propagate on *E. coli*, opening the possibility that the RnlAB and LsoAB systems act as defense mechanisms (Kai and Yonesaki, 2002; Otsuka and Yonesaki, 2012). Later it was shown that the activity of the RnlA and LsoA toxins is enhanced by RNAse HI (Naka et al., 2014) – an RNase responsible for RNA cleavage in DNA-RNA duplexes and removing the RNA primer in DNA replication (Miller et al., 1973; Itoh and Tomizawa, 1980). RNase H-fold domains are often fused to HEPN domains in larger architectures, suggesting that other RNase LS family proteins could also target RNA in DNA–RNA duplexes (Anantharaman et al., 2013). RnlA toxins in addition to the HEPN domain (also described as DBD) contain two additional domains, NTD and NRD that share similar topology of 4–5 stranded antiparallel beta-sheets with two alpha helices. It was shown that the DBD domain is responsible for dimerization of the toxin, its toxicity, and it is neutralized by both the RnlB and Dmd antitoxins (Wei et al., 2013). The functions of the NTD and NRD domains still need to be investigated.

Cas2-Like VapD Toxins

The *vapD* gene was first detected in the chromosome of the anaerobic bacterium *Dichelobacter nodosus*. The genes of this locus, prevalent in virulent strains, were designated as *vapABCD*

(Katz et al., 1992). The *vapBC* locus was later shown to encode a type II TA system (Daines et al., 2007); the same was demonstrated for *vapD* and a small upstream encoded ORF designated *vapX* (Daines et al., 2004). Later it was found that the VapD toxin from *Helicobacter pylori* displays a ferredoxin-like fold and therefore is structurally related to CRISPR-associated protein Cas2 (**Figure 3**). VapD functions as an endoribonuclease and cleaves mRNA preferentially before A or G nucleotides (Kwon et al., 2012). The presence of solitary *vapD* toxin in the *Neisseria gonorrhoeae* pEP5289 plasmid was suggested to be a factor restricting plasmids' host range. A small cryptic neisserial plasmid pJD1, however, contains full VapXD. When the VapX antitoxin from pJD1 is expressed in *E. coli*, it increases the conjugation rate of pEP5289. This suggests that solitary VapD could limit the host range of pEP5289-like plasmids to the ones that contain the pJD1 plasmid carrying an intact VapXD module, and is therefore able to neutralize incoming VapD (Pachulec and van der Does, 2010). Although it has been proposed that *vapXD* locus as well as *vapBC* locus contributes to the virulence of *H. influenzae* (Ren et al., 2012), no molecular mechanism has been demonstrated yet. VapXD satisfies the definition of the type II TA module and comprises the toxin that likely shares its origins with the Cas2 protein, however, virtually nothing else is known about its molecular mechanism or its physiological function.

mRNA CLEAVAGE IN THE RIBOSOME BY RelE TOXINS

Almost all RelE toxins cleave mRNAs in the A site of the ribosome, between the second and third position of the codon. RelE family toxins described to date share as low as 11–20% sequence identity but retain the conserved fold which is similar to ribosome independent endoribonucleases T1, Sa2, and U2 (Neubauer et al., 2009; Schureck et al., 2015). The well-studied RelE-like toxins (*E. coli* RelE, YoeB, YafQ, and *Proteus vulgaris* HigB) possess different active sites, have different preferences for targeted mRNA codons, and differ in their ability to associate with 30S and/or 70S ribosomes (Neubauer et al., 2009; Feng et al., 2013; Maehigashi et al., 2015; Schureck et al., 2016b; Pavelich et al., 2019). Analysis of the cleavage specificity of RelE homologs from a wide range of bacterial species or isolates shows that the cleavage specificity is not strict (Goeders et al., 2013). The specificity likely originates from subtle differences in the association with the ribosome, rather than recognition of specific mRNA bases. RelE toxins belong to the large superfamily of BECR-fold proteins (barnase-EndoU-colicin E5/D-RelE fold) that is found in different polymorphic toxin systems (Zhang et al., 2012).

RelE contains an antiparallel beta-sheet (usually 4-stranded) flanked by two to four surface-exposed alpha helices enriched with positively charged residues (**Figure 3**) (Neubauer et al., 2009; Schureck et al., 2016b). The positively charged residues that decorate the alpha helices and mediate interaction with the negatively charged 16S rRNA backbone are thought to be a unique feature of the ribosome-dependent RelE family of endoribonucleases (Maehigashi et al., 2015; Schureck et al., 2016b). Despite the overall structural similarity, the residues that comprise the active sites of several well-studied RelE toxins are different, as well as the ribosome conformation induced by these toxins (Neubauer et al., 2009; Feng et al., 2013; Maehigashi et al., 2015; Schureck et al., 2016b). Inside the ribosomal A site, RelE toxins reorient and activate the mRNA for 2′-OH-induced hydrolysis. Although the ribosome is not directly involved in catalysis, it is required to achieve the correct orientation of the mRNA for the cleavage reaction (Neubauer et al., 2009). Generally, RelE toxins induce strong reorganization of the mRNAs at the ribosomal A site and cause the hydrolysis between the second and third position of the codon in the A site (Neubauer et al., 2009; Feng et al., 2013; Goeders et al., 2013; Maehigashi et al., 2015; Schureck et al., 2016b). Toxins pull the mRNA out of its typical tRNA bound state. In presence of the toxin, all three A site nucleotides are shifted by more than 7 Å (Neubauer et al., 2009; Schureck et al., 2016b). Conserved residues of the toxin orient the first two A site codon bases for hydrolysis, while the third base is usually oriented with the aid of 16s rRNA (Neubauer et al., 2009; Schureck et al., 2016b). Different RelE family toxins interact with all three nucleotides of the cleaved codon in different ways, leading to subtle specificities for nucleobases at each position (Neubauer et al., 2009; Feng et al., 2013; Maehigashi et al., 2015; Schureck et al., 2016b).

The *E. coli* RelE toxin extensively relies on the ribosome for both mRNA binding and cleavage, as it has lost conserved histidine and glutamate residues used for RNA cleavage in ribosome independent RNases like RNAse T1 (Heinemann and Saenger, 1983; Takagi et al., 2005; Neubauer et al., 2009). RelE instead uses conserved basic residues both for interaction and catalysis (Neubauer et al., 2009). Almost 1/5 of the residues of RelE are basic and provide a large potential for an interaction with negatively charged RNA. In particular, a large accumulation of these residues is observed on the three helices that interact with the 16S rRNA. Other RelE toxins, such as HigB, YoeB, or YafQ do not rely on basic residues and instead function more like the RNase T1, through conserved histidine and glutamate (or histidine or tyrosine) residues for catalysis (Kamada and Hanaoka, 2005; Maehigashi et al., 2015; Schureck et al., 2016b). In addition, the RelE toxin leaves 2′-3′-cyclic phosphate at the new 3′ end (Neubauer et al., 2009), while others (YefM, YoeB, HigB) further hydrolyze it to a 3′-phosphate product (Feng et al., 2013; Maehigashi et al., 2015; Schureck et al., 2016b), like RNase T1 (Heinemann and Saenger, 1983).

Most of the RelE toxins display virtually no codon specificity and cleave between the second and third positions of the codon with the only conserved preference of purines at the third position (Goeders et al., 2013). However, several RelE family toxins were reported to be selective for the AAA lysine codon. This codon is preferred by the *E. coli* YafQ, *Rhodopseudomonas palustris* $RelE_{Rpa}$, *Nostoc* sp. $RelE_{Nsp}$, *Sinorhizobium meliloti* $RelE_{Sme}$ and *Treponema denticola* $RelE_{Tde}$ (Prysak et al., 2009; Goeders et al., 2013). *P. vulgaris* HigB cleaves A-rich codons (Hurley and Woychik, 2009; Schureck et al., 2016b), and was shown to cleave the 30S bound mRNA, indicating that it can interfere with the initiation step of translation after IF1 dissociation (Schureck et al., 2016a). The AAA and other A-rich

codons are the most frequent at the beginning of the ORFs in *E. coli* (Sato et al., 2001). Several studies have shown that the AT-rich content at the 5′ of the ORF likely reduces the secondary structures and has been shown to serve as a translation ramp for efficient protein expression in *E. coli* (Goodman et al., 2013; Verma et al., 2019). Moreover, ribosome profiling indicates that ribosomes spend a lot of time at the beginning of the transcripts (Oh et al., 2011), which might provide more time for these toxins to access their substrates. Therefore, at least a subset of RelE toxins could be considered as translation initiation inhibitors.

Several RelE toxins were shown to cleave at different sites with respect to the codon. For example, *E. coli* YhaV and *Mycobacterium avium* RelE cleave mRNA preferably between the codons (i.e., after the third base of the codon) (Goeders et al., 2013; Choi et al., 2017). However, the molecular mechanisms of substrate recognition are not yet described for these toxins. The YafO toxin was shown to be a ribosome-dependent endoribonuclease that cleaves mRNAs 11–13 bases downstream of the initiation codon (Zhang et al., 2009; Christensen-Dalsgaard et al., 2010). Such cleavage event was speculated to be located near the mRNA entrance tunnel rather than at the A site as determined for the RelE-like toxins (Schureck et al., 2016a). Although initially YafO was considered a RelE-like toxin (Christensen-Dalsgaard et al., 2010), the lack of structural information about this toxin or its interaction with ribosome lead to YafO being classified as a separate family (Leplae et al., 2011). Several RelE-family toxins, such as *E. coli* MqsR (YgiU), *Brucella abortus* BrnT and *H. pylori* HP0894 are suggested to cleave RNA in a ribosome-independent manner (Brown et al., 2009; Christensen-Dalsgaard et al., 2010; Han et al., 2011; Heaton et al., 2012) and could therefore be functionally closest to ribosome independent RNases T1, Sa2, and U2 found outside of the TA context. MqsR is a RelE-fold toxin that possesses an additional beta strand (Brown et al., 2009; Christensen-Dalsgaard et al., 2010). MqsR preferentially cleaves the 5′-G↓CN-3′ triplet (where N is preferentially U, but C or A are tolerated) in mRNA or rRNA precursors (Yamaguchi et al., 2009; Mets et al., 2017). The *H. pylori* HP0894 toxin prefers purines upstream of cleavage and its major cleavage activity was observed between first and second base at termination codons UAA and UAG (Han et al., 2011), therefore its independence from the ribosome is questionable. In fact, *E. coli* RelE was also suggested to inhibit translation termination, as in addition to CAG sense codon, it cleaves UAG or UAA stop codons between the second and third nucleotide and subsequently prevents class 1 release factors from binding the ribosome (Pedersen et al., 2003). It remains unclear whether RelE toxins targeting translation initiation would be able to compete for the A site with initiation factors, and those acting during elongation with tRNA. To date, only the affinity of YafQ toxin to the 70S assembled complex has been reported (∼360 nM) and is comparable to those of the above-mentioned factors (Maehigashi et al., 2015). It remains unclear why toxins rely on the ribosome for activity as simply blocking the translation cleavage of free mRNA would both be sufficient and efficient. It has been suggested that this dependence may indicate a specialized mechanism that would allow a response to stresses (Schureck et al., 2016a). On the other hand, selectivity for specific ribosomal complexes, such as initiating ribosomes, might be an efficient way to inhibit translation and is likely to give the same effect of reducing the global translation rate as ribosome-independent but more sequence-specific mRNA cleavage.

Bioinformatic searches and structural comparisons have detected the relationship between RelE-family toxins and those similar to ParE encoded by RK2 plasmid (Anantharaman and Aravind, 2003; Sterckx et al., 2016). ParE-family toxins act at the level of DNA replication by poisoning DNA-gyrase or by currently unidentified mechanisms (Jiang et al., 2002; Hallez et al., 2010; Yuan et al., 2010; Sterckx et al., 2016). Some residues are highly conserved across the RelE/ParE superfamily (Anantharaman and Aravind, 2003) and their three-dimensional structure is strikingly similar (Dalton and Crosson, 2010; Sterckx et al., 2016). The major differences between the mRNAses and replication inhibitors are the extended N-terminal alpha helices as well as the absence of the C-terminal helix observed in ParE toxins (Dalton and Crosson, 2010; Sterckx et al., 2016). ParE toxins are also devoid of the major catalytic residues used by RelE for mRNA cleavage in the ribosome (Neubauer et al., 2009; Dalton and Crosson, 2010). Evidence of the evolutionary relationships between RelE-like and ParE-like toxins, is the conserved principle of the binding of their cognate antitoxins to conserved hydrophobic motifs on the toxins, although different families of antitoxins can associate with both ParE and RelE-like toxins (Dalton and Crosson, 2010; Leplae et al., 2011; Sterckx et al., 2016).

INTERFERENCE WITH tRNA FUNCTIONS

VapC Toxins

The VapC toxins are characterized by the presence of a PIN domain that presents a structural similarity with the classical Rossmann-fold associated with the binding of nucleotides and nucleotide-based cofactors (Rao and Rossmann, 1973; Matelska et al., 2017; Senissar et al., 2017). The PIN domain, although originally owing its name to the type IV pili protein PilT (PilT N-terminal like nucleases), is generally found in proteins that present various endonuclease functions such as tRNA and rRNA maturation, nonsense mediated mRNA decay, and DNA replication and repair in all domains of life (Matelska et al., 2017; Senissar et al., 2017). In the PIN motif, alternating beta strands and alpha helices (α/β/α sandwich) fold into a central five stranded parallel beta-sheet decorated with alpha helices on both sides (**Figure 3**). All VapC toxins, although presenting low sequence similarity, share conserved acidic residues (D, E, D, D/N) that are distant in the primary amino acid sequence, but cluster together in the protein to form an active site that coordinates divalent-cations, such as Mg^{2+} and Mn^{2+} that are required for cleavage of single-stranded RNA (Fatica et al., 2004; Daines et al., 2007; Das et al., 2014; Matelska et al., 2017; Senissar et al., 2017). VapC toxins cleave the 3′-O-P bond of single stranded RNA to produce 3′-hydroxyl and 5′-phosphate cleavage products (McKenzie et al., 2012). So far, VapC toxins targeting the initiation $tRNA^{fMet}$, different elongation tRNAs,

and the sarcin-ricin loop (SRL) of the 23 S rRNA have been identified (**Figure 4**) (Winther and Gerdes, 2011; Winther et al., 2013, 2016). All the VapC toxins targeting tRNAs cleave in the anticodon stem-loop (ASL), either in the anticodon sequence itself or at the 3′-side of the anticodon before the stem structure (**Figure 4**). The *M. tuberculosis* VapC-mt20 cleaves the 23S rRNA SRL loop in the helix 95 that structurally mimics the ASL of tRNAs (Winther et al., 2013), therefore all VapC toxins tested so far cleave ASL-like structures. Unlike other ribonuclease toxins (namely MazF and RelE – see above), VapC toxins appear to recognize both RNA sequence and structure (Winther et al., 2013; Cruz et al., 2015; Walling and Butler, 2018; Cintron et al., 2019). Ribonucleotide modifications in the ASL region might be important for governing the specificity of VapC toxins (Winther et al., 2016; Cintron et al., 2019). Strikingly, VapC-mt20 requires the presence of the ribosome for 23S rRNA cleavage and does not cleave isolated SRL rRNA fragments (Winther et al., 2013). Since the helix 95 is exposed to the solvent at the surface of the ribosome (Yusupov et al., 2001), it is accessible for cleavage, however no information of VapC-mt20 interactions with ribosomes is currently available (Winther et al., 2013).

While some VapC toxins cleave tRNAs that service highly abundant codons, others cleave tRNAs that service rare codons, however it is unlikely that they could be compensated by other tRNAs (Cruz et al., 2015; Winther et al., 2016; Cintron et al., 2019). Therefore, it is most likely that the overexpression of VapC toxins leads to global inhibition of translation. Breaks in the ASL will interfere with the aminoacylation of those tRNAs that are recognized through their ASL by aminoacyl-tRNA synthetases (McClain et al., 1998), while charged tRNAs cleaved in ASL are unlikely to generate productive interaction with the A site of the ribosome. It has been proposed that tRNA halves produced by VapC toxins could have further functions in the stress response (Cruz et al., 2015), however the lifetime and possible interactions of VapC cleavage products remain unexplored.

Cleavage of $tRNA^{fMet}$ reported for *Salmonella*, *Shigella*, and several other VapC toxins (Winther and Gerdes, 2011) (**Figure 4**) will lead to inhibition of the initiation – a rate limiting step of protein synthesis (Laursen et al., 2005). It has been shown that expression of $VapC_{LT2}$, which leads to cleavage of initiator $tRNA^{fMet}$, would boost the translation start from elongation codons (Winther and Gerdes, 2011). However, translation of only one model mRNA was followed (Winther and Gerdes, 2011), therefore it is difficult to conclude whether induction of VapC would lead to synthesis of specific proteins initiating with other codons and whether it would lead to a sufficient, functional, and specialized proteome.

Several VapC toxins were reported to cleave the SRL loop of 23S rRNA (Winther et al., 2013; Winther et al., 2016) (**Figure 4**) that owes its name to two toxins – the α-sarcin that cleaves it at the identical position as VapC-mt20 and the ricin that removes the adenine base in the SRL (Olmo et al., 2001; Grela et al., 2019). The SRL loop is located in the GTPase associated center of the large ribosomal subunit where translational GTPases are recruited. The SRL loop is required for docking and stimulation of GTP hydrolysis of all translation GTPases (Munishkin and Wool, 1997; Rodnina et al., 1999), therefore its cleavage leads to overall inhibition of protein synthesis in prokaryotes as demonstrated for VapC, as well as in eukaryotes in case of α-sarcin (Olmo et al., 2001; Winther et al., 2013).

Numerous structures of VapCs show that these toxins are dimers and that their cognate antitoxins neutralize the toxic activity by interfering with the binding of the metal ions (Miallau et al., 2009; Dienemann et al., 2011; Min et al., 2012). However, there is still no structural information regarding their interaction with the RNA substrates, making it difficult to predict details of target recognition and catalysis. It has been demonstrated that by comparing divergent VapCs that recognize particular tRNA or rRNA substrates, they can be grouped in sub-families sharing amino acid sequence similarity (**Figure 4**). Predicted specificity of new VapCs within these families was confirmed (Winther et al., 2016). Interestingly, VapC toxins can be encoded abundantly in one genome, with an outstanding example of *M. tuberculosis* encoding as much as 50 VapC toxins (Ramage et al., 2009; Sala et al., 2014). Despite the high abundance and sequence/structure similarities, these TAs appear to not cross-talk, at least for those that are functional and that have been tested – their toxins are neutralized only by their cognate antitoxins – although in some cases this insulation could be alleviated by a single mutation (Ramage et al., 2009; Walling and Butler, 2018). Moreover, mycobacterial VapC toxins do not overlap functionally, since many of them target different tRNAs or the SRL of 23S rRNA (**Figure 4**) (Winther et al., 2016). Nevertheless, the question of the functional benefit, if any, of having large arrays of VapCs, remains open.

GNAT Toxins

GNAT-fold acetyltransferase toxins identified so far, including TacT, AtaT, and ItaT, acetylate aminocylated-tRNAs on the amino group of the cargo amino acids (Cheverton et al., 2016; Jurenas et al., 2017a; Wilcox et al., 2018). For convenience, we will refer to this family as AtaTs for Aminoacyl-tRNA-acetylating Toxins. GNAT-fold comprises a central beta-sheet composed of six to seven strands surrounded by four alpha helices (**Figure 3**). The alpha helix α3 located between the β4 and β5 strands encodes the signature motif R/Q-X-X-G-X-A/G, also referred to as 'P-loop,' which binds the pyrophosphates of acetyl-Coenzyme A which is used as a substrate for the transfer of the acetyl group (Neuwald and Landsman, 1997). The GNAT (general control non-repressible 5 (GCN5)-related *N*-acetyltransferases) family of proteins comprises more than 300,000 enzymes that acetylate various substrates from small metabolites to proteins and tRNAs (Ikeuchi et al., 2008; Salah Ud-Din et al., 2016). The GNAT-fold toxins from type II TA systems described to date form a distinct monophyletic group of GNAT acetyltransferases (Wilcox et al., 2018). Despite being related, these toxins have diverged to target different species of tRNAs charged with their respective amino acids (Jurenas et al., 2017b; Wilcox et al., 2018). The AtaT toxin encoded by the *E. coli* O157:H7 strain specifically acetylates the initiator tRNA Met-$tRNA^{fMet}$. Acetylation of the Met loaded on the Met-$tRNA^{fMet}$ impairs its interaction with the initiation factor 2 (IF2) and precludes the formation of the 30S translation initiation complex (Jurenas et al., 2017a). Formylation of the initiator fMet-$tRNA^{fMet}$ is essential for normal

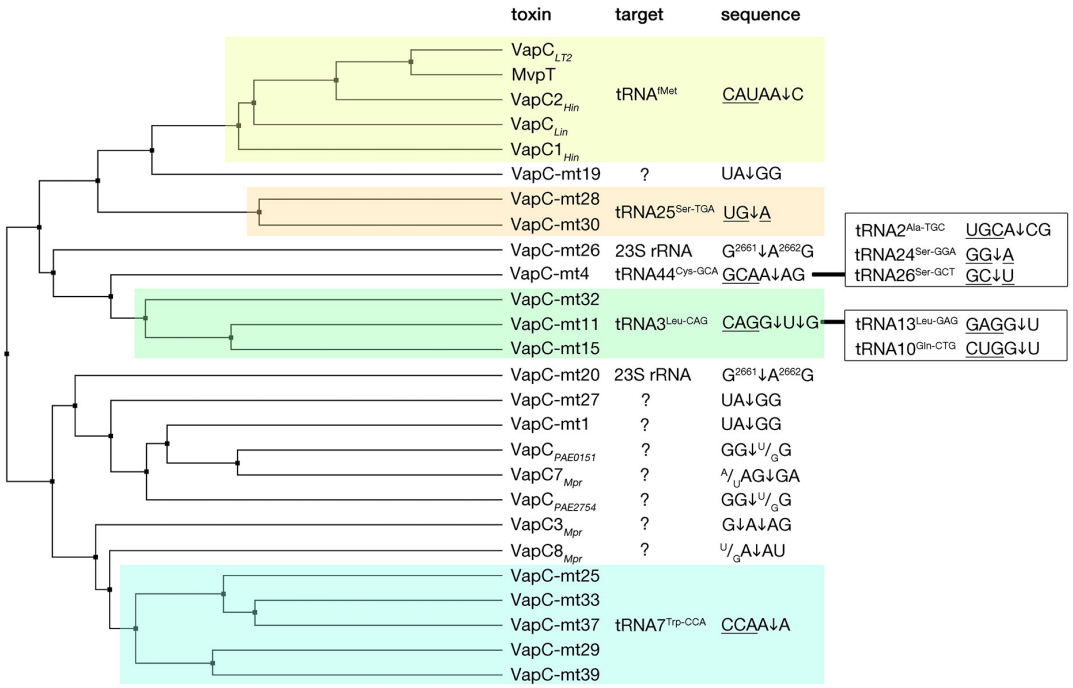

FIGURE 4 | VapC toxins and their specificity of RNA cleavage. VapC protein sequences were aligned and an average distance tree was build (BLOSUM 62, JalView) (Waterhouse et al., 2009). Substrate specificity is indicated (right). Toxin cleavage position is marked by an arrow, tRNA anticodon sequence is underlined, in the case of rRNA numbers of nucleotide positions are indicated in superscript. Additional targets and sequences are marked in the boxes on the right. Subsets of toxins sharing similar target sequences are boxed with different colors. Protein identifiers and cleavage specificity were taken from: *Haemophilus influenzae* VapC1$_{Hin}$ (NP_438487.1), VapC2$_{Hin}$ (NP_439108.1) (Walling and Butler, 2018); *Salmonella enterica* VapC$_{LT2}$ (NP_461950.1), *Shigella flexneri* MvpT (AMN61047.1) (Winther and Gerdes, 2011); *Leptospira interrogans* VapC$_{Lin}$ (AAS71220.1) (Lopes et al., 2014); *Mycobacterium tuberculosis* VapC-mt4 (NP_215109.1) (Cruz et al., 2015; Winther et al., 2016), VapC-mt11 (NP_216077.1) (Winther et al., 2016; Cintron et al., 2019), VapC-mt15(NP_216526.1), VapC-mt25 (NP_214791.1), VapC-Mt28 (NP_215123.1), VapC-mt29 (NP_215131.1), VapC-mt30(NP_215138.1), VapC-mt32(NP_215630.1), VapC-mt33 (NP_215758.1), VapC-mt37 (NP_216619.1), VapC-mt39 (NP_217046.1) (Winther et al., 2016), VapC-mt20 (NP_217065.1), VapC-mt26 (NP_215096.1) (Winther et al., 2013, 2016), VapC-mt1 (NP_214579.1) (McKenzie et al., 2012; Sharrock et al., 2018), VapC-mt19 (NP_217064.1), VapC-mt27 (NP_215112.1) (Sharrock et al., 2018), *Pyrobaculum aerophilum* VapC$_{PAE2754}$ (WP_011008882.1), VapC$_{PAE0151}$ (WP_011007068.1) (McKenzie et al., 2012); *Metallosphaera prunae* VapC3$_{Mpr}$ (WP_012020824.1), VapC7$_{Mpr}$ (WP_012021162.1), VapC8$_{Mpr}$ (WP_012021192.1) (Mukherjee et al., 2017).

growth (Shah et al., 2019) and is impaired by the Met acetylation, resulting in a strong growth inhibition and generating a dead-end product acMet-tRNA^{fMet}. *In vivo* expression of AtaT manifests in accumulation of ribosome assembly intermediates, reflecting a strong inhibition of translation initiation (Jurenas et al., 2017a). Interestingly, AtaT is able to discriminate between the initiator Met-tRNA^{fMet} and the elongator Met-tRNA^{Met} *in vitro* (Jurenas et al., 2017a) although the molecular basis of this specificity has not yet been determined. The *E. coli* HS strain encoded-ItaT toxin acetylates the elongator Ile-tRNA^{Ile}. This leads to the inhibition of translation elongation at Ile codons. It has been shown that tRNAs charged with N-blocked amino acid cannot form ternary complex EF-Tu:GTP:tRNA (Janiak et al., 1990), therefore acetylated elongator tRNAs would not be delivered to the ribosome. TacTs from *Salmonella* have been shown to have more relaxed specificities and target several elongator tRNAs. However, some specificity still occurs as different TacTs have slightly different preferences for subsets of elongation aa-tRNAs. For example, the TacT and TacT2 toxins mostly target the Gly-tRNA^{Gly}, while TacT3 prefers Ile or Leu charged tRNAs (Rycroft et al., 2018). TacTs were suggested to play an important role for persistence of *Salmonella* in macrophages (Helaine et al., 2014; Cheverton et al., 2016; Rycroft et al., 2018). However, recent data did not find any involvement of TacT or other type II TA systems from *Salmonella* in persistence (Claudi et al., 2014; Pontes and Groisman, 2019). On the other hand, like other TAs, many AtaRT-like TA systems are associated to mobile genetic elements, such as plasmids, transposons and integrons and could be involved in their maintenance (Iqbal et al., 2015; McVicker and Tang, 2016; Jurenas et al., 2017b).

HipA Toxins

HipA toxins phosphorylate aminoacyl-tRNA synthetases on conserved serines located in their ATP-binding sites, therefore leading to their inactivation (Kaspy et al., 2013; Vang Nielsen et al., 2019). Since phosphorylated aa-tRNA-synthetases cannot charge their respective tRNAs, ribosomes stall at the generated hungry codons. Consequently, an increase of ppGpp concentration is observed, associated to the activation of RelA, the effect known as the stringent response (Wendrich et al., 2002; Germain et al., 2013; Vang Nielsen et al., 2019). HipA toxins are serine-threonine kinases that belong to the phosphatidylinositol

(PI) 3/4–kinase superfamily (Correia et al., 2006). Their C-terminal domain is all-alpha helical and has a similar fold to human CDK2/cyclin A kinase (**Figure 3**) (Schumacher et al., 2009). The C-terminal domain comprises the active site with a conserved aspartate, an ATP-binding 'P-loop motif' and a Mg^{2+} binding aspartate (Correia et al., 2006). The N-terminus is an α/β globular domain specific to the HipA toxins (**Figure 3**) (Correia et al., 2006; Schumacher et al., 2009). Interestingly, it can be encoded as a separate protein, as seen in a three-component TA system HipBST from *E. coli* O27. The HipT toxin exhibits sequence similarity with the C-terminal kinase region of the HipA toxin and also targets specific tRNA synthetase (see below). The gene located upstream of the *hipT* gene, *hipS*, encodes a small protein (∼100 amino acids) corresponding to the N-terminal domain of HipA which is able to counteract the toxic activity of HipT (Vang Nielsen et al., 2019). The first gene of the operon, *hipB*, encodes a protein that enhances the ability of HipS to counteract HipT. Indeed, it has been proposed, that in addition to HipB antitoxin, the N-terminal domain of HipA (similar to the HipS protein) could be involved in regulation of HipA toxin activity through dimerization that blocks the active site (Schumacher et al., 2015).

The *E. coli* K12 HipA toxin phosphorylates the Ser 239 located in the ATP-binding site of the glutamyl-tRNA synthetase (GltX) (Germain et al., 2013; Kaspy et al., 2013). The homologous *E. coli* O127 HipT toxin phosphorylates tryptophanyl-tRNA synthetase (TrpS) at the conserved Ser 197 (corresponding to Ser 239 in the glutamyl-tRNA synthetase) (Vang Nielsen et al., 2019). The Ser 239 of GltX is located in a conserved flexible loop (characteristic to type I aa-tRNA-synthetases). Conformational changes upon $tRNA^{Glu}$ binding make this loop more exposed (Sekine et al., 2003). It was shown that HipA only phosphorylates the $tRNA^{Glu}$-bound HipA (Germain et al., 2013). The conserved motif of GltX that is phosphorylated by HipA is required for ATP binding. Thus, phosphorylation of GltX likely precludes the binding of ATP at the first step of aminoacylation reaction (Sekine et al., 2003; Germain et al., 2013). Meanwhile, the HipT toxin is able to phosphorylate the TrpS tryptophanyl-tRNA synthetase independently of $tRNA^{Trp}$ binding (Vang Nielsen et al., 2019). In fact, GltX activates glutamate to glutamyl adenylate only in presence of cognate tRNA, while TrpS can activate tryptophan to tryptophanyl-adenylate without binding $tRNA^{Trp}$ (Giege and Springer, 2016). These differences likely correspond to the conformation and accessibility of the loop in the ATP-binding site of the targeted aminoacyl-tRNA synthetases.

It has been shown that HipA inactivates itself by auto-phosphorylation (Correia et al., 2006). Typically, kinases auto-phosphorylate on the solvent exposed activation loops, HipA instead auto-phosphorylates at Ser150 located in its ATP-binding site (P-loop motif) located in the core of the protein (Schumacher et al., 2012). Autophosphorylation on Ser150 leads to conformational changes of the P-loop motif which then hinders binding of ATP (Correia et al., 2006; Schumacher et al., 2012). Likely due to the flexibility of the P-loop, autophosphorylation is an intermolecular event and therefore is likely to happen when amounts of free HipA increase, thereby providing an auto-regulation and reported as a possibility to revive from HipA-induced growth inhibition (Korch and Hill, 2006; Schumacher et al., 2012). HipT also auto-phosphorylates at Ser57 and Ser59, which are adjacent to the P-loop motif of the kinase, and which corresponds to the position of autophosphorylation of HipA (Vang Nielsen et al., 2019). The HipA toxin from *Shewanella oneidensis* was also shown to auto-phosphorylate at a similar position and it was proposed that this modification is important for complex formation with the cognate HipB antitoxin and its further binding to DNA as well as the stability of this complex (Wen et al., 2014). The autophosphorylation most likely regulates the activity and the expression of the HipA toxins.

The discovery and the name of the *E. coli K12* HipA toxin is related to the isolation of the hyper-the HipA7 persistent mutant (Moyed and Bertrand, 1983). The HipA7 strain shows a 100 to 1000-fold increase in persistence (Korch et al., 2003). It was shown that the *hipA7* allele codes for the mutations G22S and D291A in the HipA protein, which was later described to encode the toxin HipA from the *hipBA* TA system (Black et al., 1991; Korch et al., 2003). Recently, a proteomics study suggested that HipA phosphorylates multiple targets in addition to the principal target GltX, notably the ribosomal protein L11 (RplK) and other proteins involved in translation, transcription, and replication (Semanjski et al., 2018). This is in agreement with previous reports showing that HipA overexpression inhibits protein, RNA and DNA synthesis *in vivo* (Korch and Hill, 2006). However, no phosphorylation resulting from endogenous HipA encoded on the chromosome was observed, since it is repressed by HipB (Semanjski et al., 2018). Furthermore, the HipA7 strain showed phosphorylation of GltX and to a lesser extent of the phage shock protein PspA (Semanjski et al., 2018). This is in agreement with previous observations that the G22S mutation in the N-terminal domain of HipA7 likely results in compromised dimerization and failure to form a HipA:HipB:operator-DNA complex required for neutralization of HipA and transcription repression of the *hipBA* locus (Schumacher et al., 2015). Induction of HipA7 however showed more phosphorylation targets (Semanjski et al., 2018), therefore indicating that they might be physiologically non-relevant and result from a high expression of the protein. Therefore, GltX is likely to be the main target of *E. coli* K-12 HipA toxin. HipA7 also showed less auto-phosphorylation than HipA, indicating that the mutant either has less activity or compromised folding or stability, as induction of chaperons was reported upon overexpression of HipA7 (Semanjski et al., 2018). Accordingly, it has been previously reported that the HipA7 does not have a strong inhibitory effect on protein synthesis (Korch and Hill, 2006). In conclusion, the *hipA7* allele likely results in smaller effective HipA7 concentrations as compared to the wild-type system and in overall reduced activity of phosphorylation, however the HipA7 toxin is more likely to be released from the HipBA7 complex. It was later proposed that the HipBA system, together with 10 other TA systems in which the toxins are RNases, are the central effectors of antibiotic persistence in *E. coli* (Maisonneuve et al., 2018a,b; Germain et al., 2019). However, it was later demonstrated that these TA systems are not involved in persistence (see above) (Ramisetty et al., 2016; Harms et al., 2017; Shan et al., 2017;

Goormaghtigh et al., 2018a) and the idea that ppGpp is required for persistence has been challenged as well (Bhaskar et al., 2018; Pontes and Groisman, 2019).

Another family of proteins that are YjjJ-like belong to the same PI 3/4-kinase superfamily as HipA and possesses similarities in the catalytic domain (Correia et al., 2006; Maeda et al., 2017). YjjJ although encoded without cognate antitoxin was shown to be toxic, but strikingly it could be neutralized by the HipB antitoxin. YjjJ comprises a DNA-binding motif in its N-terminus that is not present in HipA-like toxins. YjjJ appears to have different cellular targets as it does not inhibit macromolecule synthesis and may affect cell division (Maeda et al., 2017).

Fic/Doc Toxins

Doc toxin is a Fic-fold protein that catalyzes phosphorylation of the elongation factor EF-Tu (Castro-Roa et al., 2013). Fic-fold proteins typically perform NMPylation (AMPylation, GMPylation, or UMPylation) as a post-translational modification on proteins using a phosphate-containing compound, usually ATP (Garcia-Pino et al., 2014; Veyron et al., 2018). In contrast, Doc toxin catalyzes the transfer of the phosphor moiety of ATP, instead of transferring the AMP, and is therefore a kinase (Castro-Roa et al., 2013). Doc phosphorylates EF-Tu at the conserved threonine Thr283 which leads to translation arrest (Castro-Roa et al., 2013). The binding site of Doc on EF-Tu likely overlaps with the tRNA binding site since ternary complex formation prevents Doc binding. In agreement with that, Doc preferentially phosphorylates the GDP-bound state of EF-Tu (Castro-Roa et al., 2013). Phosphorylation of the Thr382 located on the loop of the beta-barrel domain III of EF-Tu locks it in an unfavorable open conformation typical of GDP-bound EF-Tu (Talavera et al., 2018). Conformational dynamics of EF-Tu are the essence of its function and GTP hydrolysis has a major effect on aa-tRNA binding and interaction with the ribosome. Once locked in an open state, EF-Tu exhibits decreased affinity for aa-tRNA to a similar extent as the affinity of GDP-bound EF-Tu (Talavera et al., 2018) and is not compatible with translation (Castro-Roa et al., 2013). Fic domain toxins that perform AMPylation have also been reported to constitute type II TA modules (Harms et al., 2015). FicT toxins target TopoIV and Gyrase, and block their ATPase activity (Harms et al., 2015). Fic and Doc domain families, together sometimes referred to as Fido proteins, have conserved the alpha-helical core arranged in the bundle with two additional alpha helices perpendicular to the bundle (**Figure 3**) (Kinch et al., 2009). Although active site geometry is conserved, a single substitution in the active site motifs for Doc in comparison to Fic (K73 in Doc while G114 in Fic) leads to inverted orientation of ATP and therefore the transfer of γ-phosphate (Castro-Roa et al., 2013).

In contrast to some TA toxins, like RnlA that targets invading phages, Doc is itself encoded by a phage and provides stability to its lysogenic state. The Doc toxin is encoded by the P1 phage that is maintained as a plasmid during its lysogenic cycle. Doc was named after its impact on the remarkable stability of lysogens due to elimination of cells that have lost P1 (death on curing) (Lehnherr et al., 1993).

CONCLUSION AND DISCUSSION

Translation as a Favorite Target

Protein synthesis is one of the most complex processes in the cell. Translation involves the step by step assembly of ribosomes, coordinated movements of translation machinery at every addition of a new amino acid into the nascent polypeptide chain and well-organized termination, leading to the release of a newly synthesized protein as well as the recycling of ribosomes and translation factors (Arenz and Wilson, 2016a). The complexity of this process provides a multitude of intervention possibilities that have been explored by antibiotics, bacteriocins and secreted toxins (Zhang et al., 2012; Arenz and Wilson, 2016b; Kumariya et al., 2019). Toxins that are part of type II TA systems target translation in a multitude of ways – from destroying the transcripts before or during translation, to affecting ribosome biogenesis or interrupting the charging of tRNAs or the delivery of amino acids into growing polypeptide chains. Targeting translation allows not only choices, but also room for specialization – potential targets include a great number of tRNAs, tRNA synthetases, translation factors, and the ribosome itself (rRNAs and ribosomal proteins). Specialization of toxins portrayed in this review is already seen in almost all the TA toxin families that we know of to date. MazF toxins have diverged to target different mRNA and precursor rRNA sequences – although the majority of them cleave downstream of U and upstream of ACA nucleobases, some of them prefer an A downstream of the cleavage site and C or G upstream (**Figure 2**). Furthermore, the recently described MazF-mt9 is specialized for a particular $tRNA^{Lys43-UUU}$ species (Barth et al., 2019). VapCs also exhibit specificity for a multitude of different tRNAs or even a tRNA stem-loop structure-mimicking the 23S SRL (Winther et al., 2016) (**Figure 4**). The AtaT-like toxins that target charged tRNAs are also specific to different tRNAs and even though their toxicity relies on the acetylation of the cargo amino acid charged on its cognate tRNA, these toxins most likely recognize both the amino acids and the tRNA sequence (Jurenas et al., 2017a; Rycroft et al., 2018; Wilcox et al., 2018). Likewise, different HipA toxins have been recently demonstrated to phosphorylate different aminoacyl-tRNA synthetases (Germain et al., 2013; Kaspy et al., 2013; Vang Nielsen et al., 2019). RelE family toxins seem to recognize the mRNA pre-loaded ribosome rather than the particular mRNA sequence. Although some RelE toxins show some preference toward certain nucleobases at certain positions of a codon, the specificity comes not only from the toxin itself, but also from the conformation of ribosomes induced by the binding of these toxins (Neubauer et al., 2009; Feng et al., 2013; Maehigashi et al., 2015; Schureck et al., 2016b; Pavelich et al., 2019). Consequently, the specialization of RelEs involves the evolution of the interactions with ribosomes that in turn direct the substrate recognition.

Evolutionary Links of Type II TA Toxins

In this review, we have discussed a number of toxin families that inhibit translation at different steps. Not surprisingly, some of these families have potential evolutionary links with proteins

involved in the translation maintenance. Toxins targeting RNAs share folds with RNA metabolism associated proteins, in particular those used in RNA maturation and processing (dsRBD, HEPN and PIN-like domains) (Anantharaman and Aravind, 2006; Makarova et al., 2006; Anantharaman et al., 2013). These protein folds are not limited to type II TA toxins, they also take part in other defense and offense systems in prokaryotes as well as in eukaryotes; namely, dsRBD, HEPN, ferredoxin-like, PIN, FIC, and BECR domains can be detected in RNA interference, antitransposon and antiviral systems (Makarova et al., 2006; Kwon et al., 2012; Zhang et al., 2012; Anantharaman et al., 2013). The MazF endoribonucleases comprise a distinct SH3-like barrel-fold rather than a typical nucleic acid-binding domain. The classical SH3 domain is common in eukaryotic cell-to-cell communication and signal transduction proteins, such as signaling kinases, but is less evident in prokaryotes. It has been speculated that bacterial SH3-domain proteins act in eukaryotic cell invasion by corrupting cell signaling (Whisstock and Lesk, 1999). MazF however belongs to the group of proteins that possess domains that structurally resemble SH3, but have diverse functions and enzymatic activities (Whisstock and Lesk, 1999). Therefore, it is not clear whether the SH3 structural motif observed in MazF has a true evolutionary link with other SH3-domain proteins. Similarly, HipA toxins comprise a fold similar to (PI)3/4-kinases found in eukaryotes. These eukaryotic kinases produce 3′ phosphoinositide lipids that bind and activate proteins and therefore participate in signaling cascades (Lempiainen and Halazonetis, 2009). However, a certain class of eukaryotic PI3K family proteins are also Ser/Thr kinases (like HipA toxins) that respond to DNA damage (ATM and DNA-PKcs), nutrient stress (mTOR) or are involved in nonsense-mediated mRNA decay (SMG-1) or transcription regulation (TRRAP) (Lempiainen and Halazonetis, 2009). Whether these eukaryotic proteins have real evolutionary links with HipA toxins remains unclear.

GNAT-fold acetyltransferases are among the most abundant protein folds, however all GNAT-fold TA toxins analyzed to date target the amino group of the amino acid charged on their respective tRNAs (Cheverton et al., 2016; Jurenas et al., 2017a; Wilcox et al., 2018). GNAT-fold acetyltransferases are known to target a wide variety of substrates (Salah Ud-Din et al., 2016). A certain class of GNAT enzymes also acetylate the alpha-amino group of amino acids, however in the context of proteins, i.e., the N-terminal amino acid of peptides after methionine removal. In eukaryotes, this modification is co-translational and affects the majority of proteins, while in bacteria it is limited to specific cases of several ribosomal proteins (Vetting et al., 2008; Favrot et al., 2016). As for GNATs that interact with tRNA, TmcA – an enzyme implicated in translation fidelity – acetylates the wobble cytidine in the anticodon to prevent its misreading (Ikeuchi et al., 2008). Another family of GNATs – the Fem enzymes – use charged tRNAs as substrates for the synthesis of peptidoglycan precursors (Dare and Ibba, 2012). Fic/Doc toxins and secreted toxic effectors generally target eukaryotic and prokaryotic GTPases involved in protein signaling, translation or replication, however, new Fic targets, such as chaperons are emerging (Veyron et al., 2018). Interestingly, three TA toxin folds have also evolved to target topoisomerases (MazF/CcdB, RelE/ParE, Doc/Fic). It is not clear what the link is between translation machinery and topoisomerases. However, topoisomerases are probably the second favorite target of TA toxins. DNA damage caused by these toxins induces a SOS response and DNA repair and could favor rearrangements of genetic material providing higher chances for TAs to relocate. However, it is not clear why the mechanism of choice is blocking topoisomerases. It is worth noting that the translation machinery and topoisomerases are also among the most common targets of antibiotics. Lastly, novel enzymatic activities and targets of TA toxins have been increasingly reported, for example ADP-rybosyltransferase-fold toxins were shown to act by NAD^+ phosphorolysis and its depletion, or by DNA ribosylation (Jankevicius et al., 2016; Freire et al., 2019). These are the first examples of ADP-ribosyltransferase toxins (ART) involved in TA systems, however, many examples of secreted ART toxins are known and predicted, those also involve translation inhibitors, such as the diphtheria toxin (Zhang et al., 2012).

What Are the Roles of Type II TA Systems?

The role(s) of TA systems in bacterial physiology and evolution is a long-standing debate (Magnuson, 2007; Tsilibaris et al., 2007; Van Melderen and Saavedra De Bast, 2009; Van Melderen, 2010; Ramisetty et al., 2016; Harms et al., 2017; Culviner and Laub, 2018; Goormaghtigh et al., 2018a,b; Holden and Errington, 2018; Kaldalu et al., 2019; Mets et al., 2019; Pontes and Groisman, 2019; Wade and Laub, 2019; Fraikin et al., 2020). Since their discovery on plasmids in the 1980's and on chromosomes almost 20 years later, the TA field has been going through waves of hypothesis ranging from replicon maintenance, programmed cell death, stress response, generation of specialized ribosomes, persistence to antibiotics, to phage abortive infection mechanisms. The mainstream hypothesis for the last 10 years was the central role played by type II TA systems in persistence to antibiotics. The hypothesis is that TA systems would be induced in persister subpopulations, thereby stopping their growth and allowing these cells to tolerate the presence of antibiotics. While this hypothesis prevailed for several years, contradicting data accumulated and eventually lead to the retraction of the main papers thereby questioning the involvement of TA systems in drug tolerance (Tsilibaris et al., 2007; Conlon et al., 2016; Ramisetty et al., 2016; Harms et al., 2017; Shan et al., 2017; Goormaghtigh et al., 2018a,b; Kaldalu and Tenson, 2019; Pontes and Groisman, 2019). Although TA systems are occasionally found upregulated under stress conditions in transcriptomic data (Keren et al., 2004), this could be a natural consequence of de-repression of TA loci due to instability and degradation of the antitoxins. Since TA systems present selfish behavior, it is tempting to compare them to viral elements, and to look at TAs from the perspective of genes, and not of the organism. In this review we have provided a detailed view of the specialization of toxins sharing the same fold. Such a specialization seems to be the general trend for type II TA toxins and is seen in virtually all families that have at least several studied examples. If such systems

would be part of a stress-response system, however, one would expect the selection and conservation of the 'best' mode of action. On the contrary, the reservoir and activities of TA systems from different strains is highly variable and reminds the variety seen in offense and defense systems that are used for competition between species where innovations are beneficial. Among the driving forces for the evolution of different substrate specificities could be the tight neutralization of each toxin by its cognate antitoxin. Co-evolution of different antitoxins and different toxins relies on their vast contacts, necessary for neutralization and transcriptional autoregulation. This dependency should in addition allow for a faster evolution and selection of changes. Further, the competition between an incoming near-identical TA system likely provides selective pressure. The incoming TAs, if identical, would be neutralized by an existing copy of the TA system and such an 'anti-addiction' module therefore would prevent stable establishment of identical TAs (Van Melderen and Saavedra De Bast, 2009). Indeed, it has been shown that in some cases the different TAs in the same organism are only insulated by 1 amino acid difference (Walling and Butler, 2018), indicating that a strong selection might apply on TAs to avoid cross-talks. Lastly, the high abundance of TA systems on mobile genetic elements supports the idea that TA systems are primarily elements involved in intergenomic conflicts – inheritance of existing, defense against incoming and offense or spread of new genetic material.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

DJ was supported by a post-doctoral fellowship from the 'Fondation pour la Recherche Médicale' and EMBO non-stipendiary long-term fellowship. Work in LV's lab was supported by the Fonds National de la Recherche Scientifique (FNRS, T.0147.15F PDR and J.0061.16F CDR), the Fonds d'Encouragement à la Recherche ULB (FER-ULB), the Interuniversity Attraction Poles Program initiated by the Belgian Science Policy Office (MicroDev), the Fonds Jean Brachet, and the Fondation Van Buuren.

REFERENCES

Ahn, D. H., Lee, K. Y., Lee, S. J., Park, S. J., Yoon, H. J., Kim, S. J., et al. (2017). Structural analyses of the MazEF4 toxin-antitoxin pair in Mycobacterium tuberculosis provide evidence for a unique extracellular death factor. *J Biol Chem* 292, 18832–18847. doi: 10.1074/jbc.M117.807974

Amitai, S., Kolodkin-Gal, I., Hananya-Meltabashi, M., Sacher, A., and Engelberg-Kulka, H. (2009). *Escherichia coli* MazF leads to the simultaneous selective synthesis of both "death proteins" and "survival proteins". *PLoS Genet* 5:e1000390. doi: 10.1371/journal.pgen.1000390

Anantharaman, V., and Aravind, L. (2003). New connections in the prokaryotic toxin-antitoxin network: relationship with the eukaryotic nonsense-mediated RNA decay system. *Genome Biol* 4, R81. doi: 10.1186/gb-2003-4-12-r81

Anantharaman, V., and Aravind, L. (2006). The NYN domains: novel predicted RNAses with a PIN domain-like fold. *RNA Biol* 3, 18–27. doi: 10.4161/rna.3.1.2548

Anantharaman, V., Makarova, K. S., Burroughs, A. M., Koonin, E. V., and Aravind, L. (2013). Comprehensive analysis of the HEPN superfamily: identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing. *Biol Direct* 8, 15. doi: 10.1186/1745-6150-8-15

Arenz, S., and Wilson, D. N. (2016a). Bacterial Protein Synthesis as a Target for Antibiotic Inhibition. *Cold Spring Harb Perspect Med* 6, a025361. doi: 10.1101/cshperspect.a025361

Arenz, S., and Wilson, D. N. (2016b). Blast from the Past: Reassessing Forgotten Translation Inhibitors, Antibiotic Selectivity, and Resistance Mechanisms to Aid Drug Development. *Mol Cell* 61, 3–14. doi: 10.1016/j.molcel.2015.10.019

Barth, V. C., Zeng, J. M., Vvedenskaya, I. O., Ouyang, M., Husson, R. N., and Woychik, N. A. (2019). Toxin-mediated ribosome stalling reprograms the Mycobacterium tuberculosis proteome. *Nat Commun* 10, 3035. doi: 10.1038/s41467-019-10869-8

Belitsky, M., Avshalom, H., Erental, A., Yelin, I., Kumar, S., London, N., et al. (2011). The *Escherichia coli* extracellular death factor EDF induces the endoribonucleolytic activities of the toxins MazF and ChpBK. *Mol Cell* 41, 625–635. doi: 10.1016/j.molcel.2011.02.023

Bernard, P., and Couturier, M. (1992). Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. *J Mol Biol* 226, 735–745. doi: 10.1016/0022-2836(92)90629-x

Bernard, P., Kezdy, K. E., Van Melderen, L., Steyaert, J., Wyns, L., Pato, M. L., et al. (1993). The F plasmid CcdB protein induces efficient ATP-dependent DNA cleavage by gyrase. *J Mol Biol* 234, 534–541. doi: 10.1006/jmbi.1993.1609

Bhaskar, A., De Piano, C., Gelman, E., McKinney, J. D., and Dhar, N. (2018). Elucidating the role of (p)ppGpp in mycobacterial persistence against antibiotics. *IUBMB Life* 70, 836–844. doi: 10.1002/iub.1888

Bibi-Triki, S., Li de la Sierra-Gallay, I., Lazar, N., Leroy, A., Van Tilbeurgh, H., Sebbane, F., et al. (2014). Functional and structural analysis of HicA3-HicB3, a novel toxin-antitoxin system of Yersinia pestis. *J Bacteriol* 196, 3712–3723. doi: 10.1128/JB.01932-14

Black, D. S., Kelly, A. J., Mardis, M. J., and Moyed, H. S. (1991). Structure and organization of hip, an operon that affects lethality due to inhibition of peptidoglycan or DNA synthesis. *J Bacteriol* 173, 5732–5739. doi: 10.1128/jb.173.18.5732-5739.1991

Blower, T. R., Pei, X. Y., Short, F. L., Fineran, P. C., Humphreys, D. P., Luisi, B. F., et al. (2011). A processed noncoding RNA regulates an altruistic bacterial antiviral system. *Nat Struct Mol Biol* 18, 185–190. doi: 10.1038/nsmb.1981

Brown, B. L., Grigoriu, S., Kim, Y., Arruda, J. M., Davenport, A., Wood, T. K., et al. (2009). Three dimensional structure of the MqsR:MqsA complex: a novel TA pair comprised of a toxin homologous to RelE and an antitoxin with unique properties. *PLoS Pathog* 5:e1000706. doi: 10.1371/journal.ppat.1000706

Brown, J. M., and Shaw, K. J. (2003). A novel family of *Escherichia coli* toxin-antitoxin gene pairs. *J Bacteriol* 185, 6600–6608. doi: 10.1128/jb.185.22.6600-6608.2003

Butt, A., Higman, V. A., Williams, C., Crump, M. P., Hemsley, C. M., Harmer, N., et al. (2014). The HicA toxin from Burkholderia pseudomallei has a role in persister cell formation. *Biochem J* 459, 333–344. doi: 10.1042/BJ20140073

Castro-Roa, D., Garcia-Pino, A., De Gieter, S., van Nuland, N. A. J., Loris, R., and Zenkin, N. (2013). The Fic protein Doc uses an inverted substrate to phosphorylate and inactivate EF-Tu. *Nat Chem Biol* 9, 811–817. doi: 10.1038/nchembio.1364

Chen, R., Tu, J., Liu, Z., Meng, F., Ma, P., Ding, Z., et al. (2017). Structure of the MazF-mt9 toxin, a tRNA-specific endonuclease from Mycobacterium tuberculosis. *Biochem Biophys Res Commun* 486, 804–810. doi: 10.1016/j.bbrc.2017.03.132

Cheverton, A. M., Gollan, B., Przydacz, M., Wong, C. T., Mylona, A., Hare, S. A., et al. (2016). A *Salmonella* Toxin Promotes Persister Formation through Acetylation of tRNA. *Mol Cell* 63, 86–96. doi: 10.1016/j.molcel.2016.05.002

Choi, W., Yamaguchi, Y., Lee, J. W., Jang, K. M., Inouye, M., Kim, S. G., et al. (2017). Translation-dependent mRNA cleavage by YhaV in *Escherichia coli*. *FEBS Lett* 591, 1853–1861. doi: 10.1002/1873-3468.12705

Christensen-Dalsgaard, M., Jorgensen, M. G., and Gerdes, K. (2010). Three new RelE-homologous mRNA interferases of *Escherichia coli* differentially induced by environmental stresses. *Mol Microbiol* 75, 333–348. doi: 10.1111/j.1365-2958.2009.06969.x

Cintron, M., Zeng, J. M., Barth, V. C., Cruz, J. W., Husson, R. N., and Woychik, N. A. (2019). Accurate target identification for Mycobacterium tuberculosis endoribonuclease toxins requires expression in their native host. *Sci Rep* 9, 5949. doi: 10.1038/s41598-019-41548-9

Claudi, B., Sprote, P., Chirkova, A., Personnic, N., Zankl, J., Schurmann, N., et al. (2014). Phenotypic variation of *Salmonella* in host tissues delays eradication by antimicrobial chemotherapy. *Cell* 158, 722–733. doi: 10.1016/j.cell.2014.06.045

Conlon, B. P., Rowe, S. E., Gandt, A. B., Nuxoll, A. S., Donegan, N. P., Zalis, E. A., et al. (2016). Persister formation in Staphylococcus aureus is associated with ATP depletion. *Nat Microbiol* 1, 16051. doi: 10.1038/nmicrobiol.2016.51

Coray, D. S., Wheeler, N. E., Heinemann, J. A., and Gardner, P. P. (2017). Why so narrow: Distribution of anti-sense regulated, type I toxin-antitoxin systems compared with type II and type III systems. *RNA Biol* 14, 275–280. doi: 10.1080/15476286.2016.1272747

Correia, F. F., D'Onofrio, A., Rejtar, T., Li, L., Karger, B. L., Makarova, K., et al. (2006). Kinase activity of overexpressed HipA is required for growth arrest and multidrug tolerance in *Escherichia coli*. *J Bacteriol* 188, 8360–8367. doi: 10.1128/JB.01237-06

Cruz, J. W., Sharp, J. D., Hoffer, E. D., Maehigashi, T., Vvedenskaya, I. O., Konkimalla, A., et al. (2015). Growth-regulating Mycobacterium tuberculosis VapC-mt4 toxin is an isoacceptor-specific tRNase. *Nat Commun* 6, 7480. doi: 10.1038/ncomms8480

Culviner, P. H., and Laub, M. T. (2018). Global Analysis of the E. coli Toxin MazF Reveals Widespread Cleavage of mRNA and the Inhibition of rRNA Maturation and Ribosome Biogenesis. *Mol Cell* 70, 868.e–880.e. doi: 10.1016/j.molcel.2018.04.026

Daines, D. A., Jarisch, J., and Smith, A. L. (2004). Identification and characterization of a nontypeable *Haemophilus influenzae* putative toxin-antitoxin locus. *BMC Microbiol* 4:30. doi: 10.1186/1471-2180-4-30

Daines, D. A., Wu, M. H., and Yuan, S. Y. (2007). VapC-1 of nontypeable *Haemophilus influenzae* is a ribonuclease. *J Bacteriol* 189, 5041–5048. doi: 10.1128/JB.00290-07

Dalton, K. M., and Crosson, S. (2010). A conserved mode of protein recognition and binding in a ParD-ParE toxin-antitoxin complex. *Biochemistry* 49, 2205–2215. doi: 10.1021/bi902133s

Dao-Thi, M. H., Van Melderen, L., De Genst, E., Afif, H., Buts, L., Wyns, L., et al. (2005). Molecular basis of gyrase poisoning by the addiction toxin CcdB. *J Mol Biol* 348, 1091–1102. doi: 10.1016/j.jmb.2005.03.049

Dare, K., and Ibba, M. (2012). Roles of tRNA in cell wall biosynthesis. *Wiley Interdiscip Rev RNA* 3, 247–264. doi: 10.1002/wrna.1108

Das, U., Pogenberg, V., Subhramanyam, U. K., Wilmanns, M., Gourinath, S., and Srinivasan, A. (2014). Crystal structure of the VapBC-15 complex from Mycobacterium tuberculosis reveals a two-metal ion dependent PIN-domain ribonuclease and a variable mode of toxin-antitoxin assembly. *J Struct Biol* 188, 249–258. doi: 10.1016/j.jsb.2014.10.002

Dienemann, C., Boggild, A., Winther, K. S., Gerdes, K., and Brodersen, D. E. (2011). Crystal structure of the VapBC toxin-antitoxin complex from *Shigella* flexneri reveals a hetero-octameric DNA-binding assembly. *J Mol Biol* 414, 713–722. doi: 10.1016/j.jmb.2011.10.024

Engelberg-Kulka, H., Sat, B., Reches, M., Amitai, S., and Hazan, R. (2004). Bacterial programmed cell death systems as targets for antibiotics. *Trends Microbiol* 12, 66–71. doi: 10.1016/j.tim.2003.12.008

Fatica, A., Tollervey, D., and Dlakic, M. (2004). PIN domain of Nob1p is required for D-site cleavage in 20S pre-rRNA. *RNA* 10, 1698–1701. doi: 10.1261/rna.7123504

Favrot, L., Blanchard, J. S., and Vergnolle, O. (2016). Bacterial GCN5-Related N-Acetyltransferases: From Resistance to Regulation. *Biochemistry* 55, 989–1002. doi: 10.1021/acs.biochem.5b01269

Feng, S., Chen, Y., Kamada, K., Wang, H., Tang, K., Wang, M., et al. (2013). YoeB-ribosome structure: a canonical RNase that requires the ribosome for its specific activity. *Nucleic Acids Res* 41, 9549–9556. doi: 10.1093/nar/gkt742

Fineran, P. C., Blower, T. R., Foulds, I. J., Humphreys, D. P., Lilley, K. S., and Salmond, G. P. (2009). The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair. *Proc Natl Acad Sci U S A* 106, 894–899. doi: 10.1073/pnas.0808832106

Fraikin, N., Goormaghtigh, F., and Van Melderen, L. (2020). Type II toxin-antitoxin systems: evolution and revolutions. *J Bacteriol* 202, doi: 10.1128/JB.00763-19 **query

Fraikin, N., Rousseau, C. J., Goeders, N., and Van Melderen, L. (2019). Reassessing the Role of the Type II MqsRA Toxin-Antitoxin System in Stress Response and Biofilm Formation: mqsA Is Transcriptionally Uncoupled from mqsR. *mBio* 10, e2678–e2619. doi: 10.1128/mBio.02678-19

Freire, D. M., Gutierrez, C., Garza-Garcia, A., Grabowska, A. D., Sala, A. J., Ariyachaokun, K., et al. (2019). An NAD(+) Phosphorylase Toxin Triggers Mycobacterium tuberculosis Cell Death. *Mol Cell* 73, 1282.e–1291.e. doi: 10.1016/j.molcel.2019.01.028

Garcia-Pino, A., Balasubramanian, S., Wyns, L., Gazit, E., De Greve, H., Magnuson, R. D., et al. (2010). Allostery and intrinsic disorder mediate transcription regulation by conditional cooperativity. *Cell* 142, 101–111. doi: 10.1016/j.cell.2010.05.039

Garcia-Pino, A., Zenkin, N., and Loris, R. (2014). The many faces of Fic: structural and functional aspects of Fic enzymes. *Trends Biochem Sci* 39, 121–129. doi: 10.1016/j.tibs.2014.01.001

Gerdes, K., and Maisonneuve, E. (2012). Bacterial persistence and toxin-antitoxin loci. *Annu Rev Microbiol* 66, 103–123. doi: 10.1146/annurev-micro-092611-150159

Gerdes, K., Rasmussen, P. B., and Molin, S. (1986). Unique type of plasmid maintenance function: postsegregational killing of plasmid-free cells. *Proc Natl Acad Sci U S A* 83, 3116–3120. doi: 10.1073/pnas.83.10.3116

Gerdes, K., and Wagner, E. G. (2007). RNA antitoxins. *Curr Opin Microbiol* 10, 117–124. doi: 10.1016/j.mib.2007.03.003

Germain, E., Castro-Roa, D., Zenkin, N., and Gerdes, K. (2013). Molecular mechanism of bacterial persistence by HipA. *Mol Cell* 52, 248–254. doi: 10.1016/j.molcel.2013.08.045

Germain, E., Roghanian, M., Gerdes, K., and Maisonneuve, E. (2015). Stochastic induction of persister cells by HipA through (p)ppGpp-mediated activation of mRNA endonucleases. *Proc Natl Acad Sci U S A* 112, 5171–5176. doi: 10.1073/pnas.1423536112

Germain, E., Roghanian, M., Gerdes, K., and Maisonneuve, E. (2019). Retraction for Germain et al., Stochastic induction of persister cells by HipA through (p)ppGpp-mediated activation of mRNA endonucleases. *Proc Natl Acad Sci U S A* 116, 11077. doi: 10.1073/pnas.1906160116

Giege, R., and Springer, M. (2016). Aminoacyl-tRNA Synthetases in the Bacterial World. *EcoSal Plus* 7, doi: 10.1128/ecosalplus.ESP-0002-2016 **query*

Goeders, N., Dreze, P. L., and Van Melderen, L. (2013). Relaxed cleavage specificity within the RelE toxin family. *J Bacteriol* 195, 2541–2549. doi: 10.1128/JB.02266-12

Gogos, A., Mu, H., Bahna, F., Gomez, C. A., and Shapiro, L. (2003). Crystal structure of YdcE protein from Bacillus subtilis. *Proteins* 53, 320–322. doi: 10.1002/prot.10457

Goodman, D. B., Church, G. M., and Kosuri, S. (2013). Causes and effects of N-terminal codon bias in bacterial genes. *Science* 342, 475–479. doi: 10.1126/science.1241934

Goormaghtigh, F., Fraikin, N., Putrins, M., Hallaert, T., Hauryliuk, V., Garcia-Pino, A., et al. (2018a). Reassessing the Role of Type II Toxin-Antitoxin Systems in Formation of *Escherichia coli* Type II Persister Cells. *MBio* 9, e640–e618. doi: 10.1128/mBio.00640-18

Goormaghtigh, F., Fraikin, N., Putrins, M., Hauryliuk, V., Garcia-Pino, A., Udekwu, K., et al. (2018b). Reply to Holden and Errington, "Type II Toxin-Antitoxin Systems and Persister Cells". *MBio* 9, e1838–e1818. doi: 10.1128/mBio.01838-18

Grela, P., Szajwaj, M., Horbowicz-Drozdzal, P., and Tchorzewski, M. (2019). How Ricin Damages the Ribosome. *Toxins (Basel)* 11, 241. doi: 10.3390/toxins11050241

Guglielmini, J., and Van Melderen, L. (2011). Bacterial toxin-antitoxin systems: Translation inhibitors everywhere. *Mob Genet Elements* 1, 283–290. doi: 10.4161/mge.18477

Hallez, R., Geeraerts, D., Sterckx, Y., Mine, N., Loris, R., and Van Melderen, L. (2010). New toxins homologous to ParE belonging to three-component toxin-antitoxin systems in *Escherichia coli* O157:H7. *Mol Microbiol* 76, 719–732. doi: 10.1111/j.1365-2958.2010.07129.x

Han, K. D., Matsuura, A., Ahn, H. C., Kwon, A. R., Min, Y. H., Park, H. J., et al. (2011). Functional identification of toxin-antitoxin molecules from *Helicobacter* pylori 26695 and structural elucidation of the molecular interactions. *J Biol Chem* 286, 4842–4853. doi: 10.1074/jbc.M109.097840

Hargreaves, D., Santos-Sierra, S., Giraldo, R., Sabariegos-Jareno, R., de la Cueva-Mendez, G., Boelens, R., et al. (2002). Structural and functional analysis of the kid toxin protein from E. coli plasmid R1. *Structure* 10, 1425–1433.

Harms, A., Brodersen, D. E., Mitarai, N., and Gerdes, K. (2018). Toxins, Targets, and Triggers: An Overview of Toxin-Antitoxin Biology. *Mol Cell* 70, 768–784. doi: 10.1016/j.molcel.2018.01.003

Harms, A., Fino, C., Sorensen, M. A., Semsey, S., and Gerdes, K. (2017). Prophages and Growth Dynamics Confound Experimental Results with Antibiotic-Tolerant Persister Cells. *MBio* 8, e1964–e1917. doi: 10.1128/mBio.01964-17

Harms, A., Stanger, F. V., Scheu, P. D., de Jong, I. G., Goepfert, A., Glatter, T., et al. (2015). Adenylylation of Gyrase and Topo IV by FicT Toxins Disrupts Bacterial DNA Topology. *Cell Rep* 12, 1497–1507. doi: 10.1016/j.celrep.2015.07.056

Hayes, F., and Van Melderen, L. (2011). Toxins-antitoxins: diversity, evolution and function. *Crit Rev Biochem Mol Biol* 46, 386–408. doi: 10.3109/10409238.2011.600437

Heaton, B. E., Herrou, J., Blackwell, A. E., Wysocki, V. H., and Crosson, S. (2012). Molecular structure and function of the novel BrnT/BrnA toxin-antitoxin system of Brucella abortus. *J Biol Chem* 287, 12098–12110. doi: 10.1074/jbc.M111.332163

Heinemann, U., and Saenger, W. (1983). Crystallographic study of mechanism of ribonuclease T1-catalysed specific RNA hydrolysis. *J Biomol Struct Dyn* 1, 523–538. doi: 10.1080/07391102.1983.10507459

Helaine, S., Cheverton, A. M., Watson, K. G., Faure, L. M., Matthews, S. A., and Holden, D. W. (2014). Internalization of *Salmonella* by macrophages induces formation of nonreplicating persisters. *Science* 343, 204–208. doi: 10.1126/science.1244705

Hiraga, S., Jaffe, A., Ogura, T., Mori, H., and Takahashi, H. (1986). F plasmid ccd mechanism in *Escherichia coli*. *J Bacteriol* 166, 100–104. doi: 10.1128/jb.166.1.100-104.1986

Hoffer, E. D., Miles, S. J., and Dunham, C. M. (2017). The structure and function of Mycobacterium tuberculosis MazF-mt6 toxin provide insights into conserved features of MazF endonucleases. *J Biol Chem* 292, 7718–7726. doi: 10.1074/jbc.M117.779306

Holden, D. W., and Errington, J. (2018). Type II Toxin-Antitoxin Systems and Persister Cells. *MBio* 9, e1574–e1518. doi: 10.1128/mBio.01574-18

Huguet, K. T., Gonnet, M., Doublet, B., and Cloeckaert, A. (2016). A toxin antitoxin system promotes the maintenance of the IncA/C-mobilizable *Salmonella* Genomic Island 1. *Sci Rep* 6, 32285. doi: 10.1038/srep32285

Hurley, J. M., and Woychik, N. A. (2009). Bacterial toxin HigB associates with ribosomes and mediates translation-dependent mRNA cleavage at A-rich sites. *J Biol Chem* 284, 18605–18613. doi: 10.1074/jbc.M109.008763

Ikeuchi, Y., Kitahara, K., and Suzuki, T. (2008). The RNA acetyltransferase driven by ATP hydrolysis synthesizes N4-acetylcytidine of tRNA anticodon. *EMBO J* 27, 2194–2203. doi: 10.1038/emboj.2008.154

Iqbal, N., Guerout, A. M., Krin, E., Le Roux, F., and Mazel, D. (2015). Comprehensive Functional Analysis of the 18 *Vibrio cholerae* N16961 Toxin-Antitoxin Systems Substantiates Their Role in Stabilizing the Superintegron. *J Bacteriol* 197, 2150–2159. doi: 10.1128/JB.00108-15

Itoh, T., and Tomizawa, J. (1980). Formation of an RNA primer for initiation of replication of ColE1 DNA by ribonuclease H. *Proc Natl Acad Sci U S A* 77, 2450–2454. doi: 10.1073/pnas.77.5.2450

Jaffe, A., Ogura, T., and Hiraga, S. (1985). Effects of the ccd function of the F plasmid on bacterial growth. *J Bacteriol* 163, 841–849.

Janiak, F., Dell, V. A., Abrahamson, J. K., Watson, B. S., Miller, D. L., and Johnson, A. E. (1990). Fluorescence characterization of the interaction of various transfer RNA species with elongation factor Tu.GTP: evidence for a new functional role for elongation factor Tu in protein biosynthesis. *Biochemistry* 29, 4268–4277. doi: 10.1021/bi00470a002

Jankevicius, G., Ariza, A., Ahel, M., and Ahel, I. (2016). The Toxin-Antitoxin System DarTG Catalyzes Reversible ADP-Ribosylation of DNA. *Mol Cell* 64, 1109–1116. doi: 10.1016/j.molcel.2016.11.014

Jia, X., Yao, J., Gao, Z., Liu, G., Dong, Y. H., Wang, X., et al. (2018). Structure-function analyses reveal the molecular architecture and neutralization mechanism of a bacterial HEPN-MNT toxin-antitoxin system. *J Biol Chem* 293, 6812–6823. doi: 10.1074/jbc.RA118.002421

Jiang, Y., Pogliano, J., Helinski, D. R., and Konieczny, I. (2002). ParE toxin encoded by the broad-host-range plasmid RK2 is an inhibitor of *Escherichia coli* gyrase. *Mol Microbiol* 44, 971–979. doi: 10.1046/j.1365-2958.2002.02921.x

Jorgensen, M. G., Pandey, D. P., Jaskolska, M., and Gerdes, K. (2009). HicA of *Escherichia coli* defines a novel family of translation-independent mRNA interferases in bacteria and archaea. *J Bacteriol* 191, 1191–1199. doi: 10.1128/JB.01013-08

Jurenas, D., Chatterjee, S., Konijnenberg, A., Sobott, F., Droogmans, L., Garcia-Pino, A., et al. (2017a). AtaT blocks translation initiation by N-acetylation of the initiator tRNA(fMet). *Nat Chem Biol* 13, 640–646. doi: 10.1038/nchembio.2346

Jurenas, D., Garcia-Pino, A., and Van Melderen, L. (2017b). Novel toxins from type II toxin-antitoxin systems with acetyltransferase activity. *Plasmid* 93, 30–35. doi: 10.1016/j.plasmid.2017.08.005

Jurenas, D., Van Melderen, L., and Garcia-Pino, A. (2019). Mechanism of regulation and neutralization of the AtaR-AtaT toxin-antitoxin system. *Nat Chem Biol* 15, 285–294. doi: 10.1038/s41589-018-0216-z

Kai, T., and Yonesaki, T. (2002). Multiple mechanisms for degradation of bacteriophage T4 soc mRNA. *Genetics* 160, 5–12.

Kaldalu, N., Maivali, U., Hauryliuk, V., and Tenson, T. (2019). Reanalysis of Proteomics Results Fails To Detect MazF-Mediated Stress Proteins. *MBio* 10, e949–e919. doi: 10.1128/mBio.00949-19

Kaldalu, N., and Tenson, T. (2019). Slow growth causes bacterial persistence. *Sci Signal* 12, eaay1167. doi: 10.1126/scisignal.aay1167

Kamada, K., and Hanaoka, F. (2005). Conformational change in the catalytic site of the ribonuclease YoeB toxin by YefM antitoxin. *Mol Cell* 19, 497–509. doi: 10.1016/j.molcel.2005.07.004

Kamada, K., Hanaoka, F., and Burley, S. K. (2003). Crystal structure of the MazE/MazF complex: molecular bases of antidote-toxin recognition. *Mol Cell* 11, 875–884.

Karoui, H., Bex, F., Dreze, P., and Couturier, M. (1983). Ham22, a mini-F mutation which is lethal to host cell and promotes recA-dependent induction of lambdoid prophage. *EMBO J* 2, 1863–1868.

Kaspy, I., Rotem, E., Weiss, N., Ronin, I., Balaban, N. Q., and Glaser, G. (2013). HipA-mediated antibiotic persistence via phosphorylation of the glutamyl-tRNA-synthetase. *Nat Commun* 4, 3001. doi: 10.1038/ncomms4001

Katz, M. E., Strugnell, R. A., and Rood, J. I. (1992). Molecular characterization of a genomic region associated with virulence in Dichelobacter nodosus. *Infect Immun* 60, 4586–4592.

Keren, I., Shah, D., Spoering, A., Kaldalu, N., and Lewis, K. (2004). Specialized persister cells and the mechanism of multidrug tolerance in *Escherichia coli*. *J Bacteriol* 186, 8172–8180. doi: 10.1128/JB.186.24.8172-8180.2004

Kim, D. H., Kang, S. M., Park, S. J., Jin, C., Yoon, H. J., and Lee, B. J. (2018). Functional insights into the Streptococcus pneumoniae HicBA toxin-antitoxin system based on a structural study. *Nucleic Acids Res* 46, 6371–6386. doi: 10.1093/nar/gky469

Kinch, L. N., Yarbrough, M. L., Orth, K., and Grishin, N. V. (2009). Fido, a novel AMPylation domain common to fic, doc, and AvrB. *PLoS One* 4:e5818. doi: 10.1371/journal.pone.0005818

Koga, M., Otsuka, Y., Lemire, S., and Yonesaki, T. (2011). *Escherichia coli* rnlA and rnlB compose a novel toxin-antitoxin system. *Genetics* 187, 123–130. doi: 10.1534/genetics.110.121798

Kolodkin-Gal, I., Hazan, R., Gaathon, A., Carmeli, S., and Engelberg-Kulka, H. (2007). A linear pentapeptide is a quorum-sensing factor required for mazEF-mediated cell death in *Escherichia coli*. *Science* 318, 652–655. doi: 10.1126/science.1147248

Korch, S. B., Henderson, T. A., and Hill, T. M. (2003). Characterization of the hipA7 allele of *Escherichia coli* and evidence that high persistence is governed by (p)ppGpp synthesis. *Mol Microbiol* 50, 1199–1213. doi: 10.1046/j.1365-2958.2003.03779.x

Korch, S. B., and Hill, T. M. (2006). Ectopic overexpression of wild-type and mutant hipA genes in *Escherichia coli*: effects on macromolecular synthesis and persister formation. *J Bacteriol* 188, 3826–3836. doi: 10.1128/JB.01740-05

Kumar, S., Kolodkin-Gal, I., and Engelberg-Kulka, H. (2013). Novel quorum-sensing peptides mediating interspecies bacterial cell death. *MBio* 4, 1–12. doi: 10.1128/mBio.00314-13

Kumariya, R., Garsa, A. K., Rajput, Y. S., Sood, S. K., Akhtar, N., and Patel, S. (2019). Bacteriocins: Classification, synthesis, mechanism of action and resistance development in food spoilage causing bacteria. *Microb Pathog* 128, 171–177. doi: 10.1016/j.micpath.2019.01.002

Kwon, A. R., Kim, J. H., Park, S. J., Lee, K. Y., Min, Y. H., Im, H., et al. (2012). Structural and biochemical characterization of HP0315 from *Helicobacter* pylori as a VapD protein with an endoribonuclease activity. *Nucleic Acids Res* 40, 4216–4228. doi: 10.1093/nar/gkr1305

Laursen, B. S., Sorensen, H. P., Mortensen, K. K., and Sperling-Petersen, H. U. (2005). Initiation of protein synthesis in bacteria. *Microbiol Mol Biol Rev* 69, 101–123. doi: 10.1128/MMBR.69.1.101-123.2005

Lehnherr, H., Maguin, E., Jafri, S., and Yarmolinsky, M. B. (1993). Plasmid addiction genes of bacteriophage P1: doc, which causes cell death on curing of prophage, and phd, which prevents host death when prophage is retained. *J Mol Biol* 233, 414–428. doi: 10.1006/jmbi.1993.1521

Lempiainen, H., and Halazonetis, T. D. (2009). Emerging common themes in regulation of PIKKs and PI3Ks. *EMBO J* 28, 3067–3073. doi: 10.1038/emboj.2009.281

Leplae, R., Geeraerts, D., Hallez, R., Guglielmini, J., Dreze, P., and Van Melderen, L. (2011). Diversity of bacterial type II toxin-antitoxin systems: a comprehensive search and functional analysis of novel families. *Nucleic Acids Res* 39, 5513–5525. doi: 10.1093/nar/gkr131

Li, G. Y., Zhang, Y., Inouye, M., and Ikura, M. (2009). Inhibitory mechanism of *Escherichia coli* RelE-RelB toxin-antitoxin module involves a helix displacement near an mRNA interferase active site. *J Biol Chem* 284, 14628–14636. doi: 10.1074/jbc.M809656200

Lopes, A. P., Lopes, L. M., Fraga, T. R., Chura-Chambi, R. M., Sanson, A. L., Cheng, E., et al. (2014). VapC from the leptospiral VapBC toxin-antitoxin module displays ribonuclease activity on the initiator tRNA. *PLoS One* 9:e101678. doi: 10.1371/journal.pone.0101678

Loris, R., Dao-Thi, M. H., Bahassi, E. M., Van Melderen, L., Poortmans, F., Liddington, R., et al. (1999). Crystal structure of CcdB, a topoisomerase poison from E. coli. *J Mol Biol* 285, 1667–1677. doi: 10.1006/jmbi.1998.2395

Maeda, Y., Lin, C. Y., Ishida, Y., Inouye, M., Yamaguchi, Y., and Phadtare, S. (2017). Characterization of YjjJ toxin of *Escherichia coli*. *FEMS Microbiol Lett* 364, fnx086. doi: 10.1093/femsle/fnx086

Maehigashi, T., Ruangprasert, A., Miles, S. J., and Dunham, C. M. (2015). Molecular basis of ribosome recognition and mRNA hydrolysis by the E. coli YafQ toxin. *Nucleic Acids Res* 43, 8002–8012. doi: 10.1093/nar/gkv791

Magnuson, R. D. (2007). Hypothetical functions of toxin-antitoxin systems. *J Bacteriol* 189, 6089–6092. doi: 10.1128/JB.00958-07

Maisonneuve, E., Castro-Camargo, M., and Gerdes, K. (2013). (p)ppGpp controls bacterial persistence by stochastic induction of toxin-antitoxin activity. *Cell* 154, 1140–1150. doi: 10.1016/j.cell.2013.07.048

Maisonneuve, E., Castro-Camargo, M., and Gerdes, K. (2018a). (p)ppGpp Controls Bacterial Persistence by Stochastic Induction of Toxin-Antitoxin Activity. *Cell* 172, 1135. doi: 10.1016/j.cell.2018.02.023

Maisonneuve, E., Shakespeare, L. J., Jorgensen, M. G., and Gerdes, K. (2011). Bacterial persistence by RNA endonucleases. *Proc Natl Acad Sci U S A* 108, 13206–13211. doi: 10.1073/pnas.1100186108

Maisonneuve, E., Shakespeare, L. J., Jorgensen, M. G., and Gerdes, K. (2018b). Retraction for Maisonneuve et al., Bacterial persistence by RNA endonucleases. *Proc Natl Acad Sci U S A* 115, E2901. doi: 10.1073/pnas.1803278115

Makarova, K. S., Grishin, N. V., and Koonin, E. V. (2006). The HicAB cassette, a putative novel, RNA-targeting toxin-antitoxin system in archaea and bacteria. *Bioinformatics* 22, 2581–2584. doi: 10.1093/bioinformatics/btl418

Manav, M. C., Turnbull, K. J., Jurenas, D., Garcia-Pino, A., Gerdes, K., and Brodersen, D. E. (2019). The E. coli HicB Antitoxin Contains a Structurally Stable Helix-Turn-Helix DNA Binding Domain. *Structure* 27, 1675.e–1685.e. doi: 10.1016/j.str.2019.08.008

Matelska, D., Steczkiewicz, K., and Ginalski, K. (2017). Comprehensive classification of the PIN domain-like superfamily. *Nucleic Acids Res* 45, 6995–7020. doi: 10.1093/nar/gkx494

McClain, W. H., Schneider, J., Bhattacharya, S., and Gabriel, K. (1998). The importance of tRNA backbone-mediated interactions with synthetase for aminoacylation. *Proc Natl Acad Sci U S A* 95, 460–465. doi: 10.1073/pnas.95.2.460

McKenzie, J. L., Duyvestyn, J. M., Smith, T., Bendak, K., Mackay, J., Cursons, R., et al. (2012). Determination of ribonuclease sequence-specificity using Pentaprobes and mass spectrometry. *RNA* 18, 1267–1278. doi: 10.1261/rna.031229.111

McVicker, G., and Tang, C. M. (2016). Deletion of toxin-antitoxin systems in the evolution of *Shigella* sonnei as a host-adapted pathogen. *Nat Microbiol* 2, 16204. doi: 10.1038/nmicrobiol.2016.204

Mets, T., Kasvandik, S., Saarma, M., Maivali, U., Tenson, T., and Kaldalu, N. (2019). Fragmentation of *Escherichia coli* mRNA by MazF and MqsR. *Biochimie* 156, 79–91. doi: 10.1016/j.biochi.2018.10.004

Mets, T., Lippus, M., Schryer, D., Liiv, A., Kasari, V., Paier, A., et al. (2017). Toxins MazF and MqsR cleave *Escherichia coli* rRNA precursors at multiple sites. *RNA Biol* 14, 124–135. doi: 10.1080/15476286.2016.1259784

Mhlanga-Mutangadura, T., Morlin, G., Smith, A. L., Eisenstark, A., and Golomb, M. (1998). Evolution of the major pilus gene cluster of *Haemophilus influenzae*. *J Bacteriol* 180, 4693–4703.

Miallau, L., Faller, M., Chiang, J., Arbing, M., Guo, F., Cascio, D., et al. (2009). Structure and proposed activity of a member of the VapBC family of toxin-antitoxin systems. VapBC-5 from Mycobacterium tuberculosis. *J Biol Chem* 284, 276–283. doi: 10.1074/jbc.M805061200

Miller, H. I., Riggs, A. D., and Gill, G. N. (1973). Ribonuclease H (hybrid) in *Escherichia coli*. Identification and characterization. *J Biol Chem* 248, 2621–2624.

Min, A. B., Miallau, L., Sawaya, M. R., Habel, J., Cascio, D., and Eisenberg, D. (2012). The crystal structure of the Rv0301-Rv0300 VapBC-3 toxin-antitoxin complex from M. tuberculosis reveals a Mg(2)(+) ion in the active site and a putative RNA-binding site. *Protein Sci* 21, 1754–1767. doi: 10.1002/pro.2161

Miyamoto, T., Kato, Y., Sekiguchi, Y., Tsuneda, S., and Noda, N. (2016a). Characterization of MazF-Mediated Sequence-Specific RNA Cleavage in *Pseudomonas* putida Using Massive Parallel Sequencing. *PLoS One* 11:e0149494. doi: 10.1371/journal.pone.0149494

Miyamoto, T., Ota, Y., Yokota, A., Suyama, T., Tsuneda, S., and Noda, N. (2017). Characterization of a Deinococcus radiodurans MazF: A UACA-specific RNA endoribonuclease. *Microbiologyopen* 6, e00501. doi: 10.1002/mbo3.501

Miyamoto, T., Yokota, A., Ota, Y., Tsuruga, M., Aoi, R., Tsuneda, S., et al. (2018). Nitrosomonas europaea MazF Specifically Recognises the UGG Motif and Promotes Selective RNA Degradation. *Front Microbiol* 9:2386. doi: 10.3389/fmicb.2018.02386

Miyamoto, T., Yokota, A., Tsuneda, S., and Noda, N. (2016b). AAU-Specific RNA Cleavage Mediated by MazF Toxin Endoribonuclease Conserved in Nitrosomonas europaea. *Toxins (Basel)* 8, 174. doi: 10.3390/toxins8060174

Moyed, H. S., and Bertrand, K. P. (1983). hipA, a newly recognized gene of *Escherichia coli* K-12 that affects frequency of persistence after inhibition of murein synthesis. *J Bacteriol* 155, 768–775.

Mukherjee, A., Wheaton, G. H., Counts, J. A., Ijeomah, B., Desai, J., and Kelly, R. M. (2017). VapC toxins drive cellular dormancy under uranium stress for the extreme thermoacidophile Metallosphaera prunae. *Environ Microbiol* 19, 2831–2842. doi: 10.1111/1462-2920.13808

Munishkin, A., and Wool, I. G. (1997). The ribosome-in-pieces: binding of elongation factor EF-G to oligoribonucleotides that mimic the sarcin/ricin and thiostrepton domains of 23S ribosomal RNA. *Proc Natl Acad Sci U S A* 94, 12280–12284. doi: 10.1073/pnas.94.23.12280

Munoz-Gomez, A. J., Santos-Sierra, S., Berzal-Herranz, A., Lemonnier, M., and Diaz-Orejas, R. (2004). Insights into the specificity of RNA cleavage by the *Escherichia coli* MazF toxin. *FEBS Lett* 567, 316–320. doi: 10.1016/j.febslet.2004.05.005

Naka, K., Koga, M., Yonesaki, T., and Otsuka, Y. (2014). RNase HI stimulates the activity of RnlA toxin in *Escherichia coli*. *Mol Microbiol* 91, 596–605. doi: 10.1111/mmi.12479

Nariya, H., and Inouye, M. (2008). MazF, an mRNA interferase, mediates programmed cell death during multicellular Myxococcus development. *Cell* 132, 55–66. doi: 10.1016/j.cell.2007.11.044

Neubauer, C., Gao, Y. G., Andersen, K. R., Dunham, C. M., Kelley, A. C., Hentschel, J., et al. (2009). The structural basis for mRNA recognition and cleavage by the ribosome-dependent endonuclease RelE. *Cell* 139, 1084–1095. doi: 10.1016/j.cell.2009.11.015

Neuwald, A. F., and Landsman, D. (1997). GCN5-related histone N-acetyltransferases belong to a diverse superfamily that includes the yeast SPT10 protein. *Trends Biochem Sci* 22, 154–155.

Ogura, T., and Hiraga, S. (1983). Partition mechanism of F plasmid: two plasmid gene-encoded products and a cis-acting region are involved in partition. *Cell* 32, 351–360. doi: 10.1016/0092-8674(83)90454-3

Oh, E., Becker, A. H., Sandikci, A., Huber, D., Chaba, R., Gloge, F., et al. (2011). Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. *Cell* 147, 1295–1308. doi: 10.1016/j.cell.2011.10.044

Olmo, N., Turnay, J., Gonzalez de Buitrago, G., Lopez, de Silanes, I., Gavilanes, J. G., et al. (2001). Cytotoxic mechanism of the ribotoxin alpha-sarcin. Induction of cell death via apoptosis. *Eur J Biochem* 268, 2113–2123. doi: 10.1046/j.1432-1327.2001.02086.x

Otsuka, Y., Koga, M., Iwamoto, A., and Yonesaki, T. (2007). A role of RnlA in the RNase LS activity from *Escherichia coli*. *Genes Genet Syst* 82, 291–299.

Otsuka, Y., and Yonesaki, T. (2012). Dmd of bacteriophage T4 functions as an antitoxin against *Escherichia coli* LsoA and RnlA toxins. *Mol Microbiol* 83, 669–681. doi: 10.1111/j.1365-2958.2012.07975.x

Pachulec, E., and van der Does, C. (2010). Conjugative plasmids of *Neisseria gonorrhoeae*. *PLoS One* 5:e9962. doi: 10.1371/journal.pone.0009962

Page, R., and Peti, W. (2016). Toxin-antitoxin systems in bacterial growth arrest and persistence. *Nat Chem Biol* 12, 208–214. doi: 10.1038/nchembio.2044

Pandey, D. P., and Gerdes, K. (2005). Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Res* 33, 966–976. doi: 10.1093/nar/gki201

Park, J. H., Yamaguchi, Y., and Inouye, M. (2011). Bacillus subtilis MazF-bs (EndoA) is a UACAU-specific mRNA interferase. *FEBS Lett* 585, 2526–2532. doi: 10.1016/j.febslet.2011.07.008

Pavelich, I. J., Maehigashi, T., Hoffer, E. D., Ruangprasert, A., Miles, S. J., and Dunham, C. M. (2019). Monomeric YoeB toxin retains RNase activity but adopts an obligate dimeric form for thermal stability. *Nucleic Acids Res* 47, 10400–10413. doi: 10.1093/nar/gkz760

Pedersen, K., Zavialov, A. V., Pavlov, M. Y., Elf, J., Gerdes, K., and Ehrenberg, M. (2003). The bacterial toxin RelE displays codon-specific cleavage of mRNAs in the ribosomal A site. *Cell* 112, 131–140. doi: 10.1016/s0092-8674(02)01248-5

Pontes, M. H., and Groisman, E. A. (2019). Slow growth determines nonheritable antibiotic resistance in *Salmonella enterica*. *Sci Signal* 12, eaax3938. doi: 10.1126/scisignal.aax3938

Prysak, M. H., Mozdzierz, C. J., Cook, A. M., Zhu, L., Zhang, Y., Inouye, M., et al. (2009). Bacterial toxin YafQ is an endoribonuclease that associates with the ribosome and blocks translation elongation through sequence-specific and frame-dependent mRNA cleavage. *Mol Microbiol* 71, 1071–1087. doi: 10.1111/j.1365-2958.2008.06572.x

Ramage, H. R., Connolly, L. E., and Cox, J. S. (2009). Comprehensive functional analysis of Mycobacterium tuberculosis toxin-antitoxin systems: implications for pathogenesis, stress responses, and evolution. *PLoS Genet* 5:e1000767. doi: 10.1371/journal.pgen.1000767

Ramisetty, B. C., Ghosh, D., Roy Chowdhury, M., and Santhosh, R. S. (2016). What Is the Link between Stringent Response, Endoribonuclease Encoding Type II Toxin-Antitoxin Systems and Persistence? *Front Microbiol* 7:1882. doi: 10.3389/fmicb.2016.01882

Rao, S. T., and Rossmann, M. G. (1973). Comparison of super-secondary structures in proteins. *J Mol Biol* 76, 241–256. doi: 10.1016/0022-2836(73)90388-4

Ren, D., Walker, A. N., and Daines, D. A. (2012). Toxin-antitoxin loci vapBC-1 and vapXD contribute to survival and virulence in nontypeable *Haemophilus influenzae*. *BMC Microbiol* 12:263. doi: 10.1186/1471-2180-12-263

Rodnina, M. V., Savelsbergh, A., and Wintermeyer, W. (1999). Dynamics of translation on the ribosome: molecular mechanics of translocation. *FEMS Microbiol Rev* 23, 317–333. doi: 10.1111/j.1574-6976.1999.tb00402.x

Rothenbacher, F. P., Suzuki, M., Hurley, J. M., Montville, T. J., Kirn, T. J., Ouyang, M., et al. (2012). Clostridium difficile MazF toxin exhibits selective, not global, mRNA cleavage. *J Bacteriol* 194, 3464–3474. doi: 10.1128/JB.00217-12

Rycroft, J. A., Gollan, B., Grabe, G. J., Hall, A., Cheverton, A. M., Larrouy-Maumus, G., et al. (2018). Activity of acetyltransferase toxins involved in *Salmonella* persister formation during macrophage infection. *Nat Commun* 9, 1993. doi: 10.1038/s41467-018-04472-6

Ryter, J. M., and Schultz, S. C. (1998). Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. *EMBO J* 17, 7505–7513. doi: 10.1093/emboj/17.24.7505

Sala, A., Bordes, P., and Genevaux, P. (2014). Multiple toxin-antitoxin systems in Mycobacterium tuberculosis. *Toxins (Basel)* 6, 1002–1020. doi: 10.3390/toxins6031002

Salah Ud-Din, A. I., Tikhomirova, A., and Roujeinikova, A. (2016). Structure and Functional Diversity of GCN5-Related N-Acetyltransferases (GNAT). *Int J Mol Sci* 17, 1018. doi: 10.3390/ijms17071018

Sato, T., Terabe, M., Watanabe, H., Gojobori, T., Hori-Takemoto, C., and Miura, K. (2001). Codon and base biases after the initiation codon of the open reading frames in the *Escherichia coli* genome and their influence on the translation efficiency. *J Biochem* 129, 851–860. doi: 10.1093/oxfordjournals.jbchem.a002929

Sauert, M., Wolfinger, M. T., Vesper, O., Muller, C., Byrgazov, K., and Moll, I. (2016). The MazF-regulon: a toolbox for the post-transcriptional stress response in *Escherichia coli*. *Nucleic Acids Res* 44, 6660–6675. doi: 10.1093/nar/gkw115

Schifano, J. M., Cruz, J. W., Vvedenskaya, I. O., Edifor, R., Ouyang, M., Husson, R. N., et al. (2016). tRNA is a new target for cleavage by a MazF toxin. *Nucleic Acids Res* 44, 1256–1270. doi: 10.1093/nar/gkv1370

Schifano, J. M., Edifor, R., Sharp, J. D., Ouyang, M., Konkimalla, A., Husson, R. N., et al. (2013). Mycobacterial toxin MazF-mt6 inhibits translation through cleavage of 23S rRNA at the ribosomal A site. *Proc Natl Acad Sci U S A* 110, 8501–8506. doi: 10.1073/pnas.1222031110

Schifano, J. M., Vvedenskaya, I. O., Knoblauch, J. G., Ouyang, M., Nickels, B. E., and Woychik, N. A. (2014). An RNA-seq method for defining endoribonuclease cleavage specificity identifies dual rRNA substrates for toxin MazF-mt3. *Nat Commun* 5, 3538. doi: 10.1038/ncomms4538

Schumacher, M. A., Balani, P., Min, J., Chinnam, N. B., Hansen, S., Vulic, M., et al. (2015). HipBA-promoter structures reveal the basis of heritable multidrug tolerance. *Nature* 524, 59–64. doi: 10.1038/nature14662

Schumacher, M. A., Min, J., Link, T. M., Guan, Z., Xu, W., Ahn, Y. H., et al. (2012). Role of unusual P loop ejection and autophosphorylation in HipA-mediated persistence and multidrug tolerance. *Cell Rep* 2, 518–525. doi: 10.1016/j.celrep.2012.08.013

Schumacher, M. A., Piro, K. M., Xu, W., Hansen, S., Lewis, K., and Brennan, R. G. (2009). Molecular mechanisms of HipA-mediated multidrug tolerance and its neutralization by HipB. *Science* 323, 396–401. doi: 10.1126/science.1163806

Schureck, M. A., Dunkle, J. A., Maehigashi, T., Miles, S. J., and Dunham, C. M. (2015). Defining the mRNA recognition signature of a bacterial toxin protein. *Proc Natl Acad Sci U S A* 112, 13862–13867. doi: 10.1073/pnas.1512959112

Schureck, M. A., Maehigashi, T., Miles, S. J., Marquez, J., and Dunham, C. M. (2016a). mRNA bound to the 30S subunit is a HigB toxin substrate. *RNA* 22, 1261–1270. doi: 10.1261/rna.056218.116

Schureck, M. A., Repack, A., Miles, S. J., Marquez, J., and Dunham, C. M. (2016b). Mechanism of endonuclease cleavage by the HigB toxin. *Nucleic Acids Res* 44, 7944–7953. doi: 10.1093/nar/gkw598

Schuster, C. F., Park, J. H., Prax, M., Herbig, A., Nieselt, K., Rosenstein, R., et al. (2013). Characterization of a mazEF toxin-antitoxin homologue from Staphylococcus equorum. *J Bacteriol* 195, 115–125. doi: 10.1128/JB.00400-12

Sekine, S., Nureki, O., Dubois, D. Y., Bernier, S., Chenevert, R., Lapointe, J., et al. (2003). ATP binding by glutamyl-tRNA synthetase is switched to the productive mode by tRNA binding. *EMBO J* 22, 676–688. doi: 10.1093/emboj/cdg053

Semanjski, M., Germain, E., Bratl, K., Kiessling, A., Gerdes, K., and Macek, B. (2018). The kinases HipA and HipA7 phosphorylate different substrate pools in *Escherichia coli* to promote multidrug tolerance. *Sci Signal* 11, eaat5750. doi: 10.1126/scisignal.aat5750

Senissar, M., Manav, M. C., and Brodersen, D. E. (2017). Structural conservation of the PIN domain active site across all domains of life. *Protein Sci* 26, 1474–1492. doi: 10.1002/pro.3193

Shah, R. A., Varada, R., Sah, S., Shetty, S., Lahry, K., Singh, S., et al. (2019). Rapid formylation of the cellular initiator tRNA population makes a crucial contribution to its exclusive participation at the step of initiation. *Nucleic Acids Res* 47, 1908–1919. doi: 10.1093/nar/gky1310

Shaku, M., Park, J. H., Inouye, M., and Yamaguchi, Y. (2018). Identification of MazF Homologue in *Legionella pneumophila* Which Cleaves RNA at the AACU Sequence. *J Mol Microbiol Biotechnol* 28, 269–280. doi: 10.1159/000497146

Shan, Y., Brown Gandt, A., Rowe, S. E., Deisinger, J. P., Conlon, B. P., and Lewis, K. (2017). ATP-Dependent Persister Formation in *Escherichia coli*. *MBio* 8, e2267–e2216. doi: 10.1128/mBio.02267-16

Sharrock, A., Ruthe, A., Andrews, E. S. V., Arcus, V. A., and Hicks, J. L. (2018). VapC proteins from Mycobacterium tuberculosis share ribonuclease sequence specificity but differ in regulation and toxicity. *PLoS One* 13:e0203412. doi: 10.1371/journal.pone.0203412

Simanshu, D. K., Yamaguchi, Y., Park, J. H., Inouye, M., and Patel, D. J. (2013). Structural basis of mRNA recognition and cleavage by toxin MazF and its regulation by antitoxin MazE in Bacillus subtilis. *Mol Cell* 52, 447–458. doi: 10.1016/j.molcel.2013.09.006

Sterckx, Y. G., Jove, T., Shkumatov, A. V., Garcia-Pino, A., Geerts, L., De Kerpel, M., et al. (2016). A unique hetero-hexadecameric architecture displayed by the *Escherichia coli* O157 PaaA2-ParE2 antitoxin-toxin complex. *J Mol Biol* 428, 1589–1603. doi: 10.1016/j.jmb.2016.03.007

Szekeres, S., Dauti, M., Wilde, C., Mazel, D., and Rowe-Magnus, D. A. (2007). Chromosomal toxin-antitoxin loci can diminish large-scale genome reductions in the absence of selection. *Mol Microbiol* 63, 1588–1605. doi: 10.1111/j.1365-2958.2007.05613.x

Takagi, H., Kakuta, Y., Okada, T., Yao, M., Tanaka, I., and Kimura, M. (2005). Crystal structure of archaeal toxin-antitoxin RelE-RelB complex with implications for toxin activity and antitoxin effects. *Nat Struct Mol Biol* 12, 327–331. doi: 10.1038/nsmb911

Talavera, A., Hendrix, J., Versees, W., Jurenas, D., Van Nerom, K., Vandenberk, N., et al. (2018). Phosphorylation decelerates conformational dynamics in bacterial translation elongation factors. *Sci Adv* 4, eaa9714. doi: 10.1126/sciadv.aap9714

Tam, J. E., and Kline, B. C. (1989). The F plasmid ccd autorepressor is a complex of CcdA and CcdB proteins. *Mol Gen Genet* 219, 26–32. doi: 10.1007/bf00261153

Thomet, M., Trautwetter, A., Ermel, G., and Blanco, C. (2019). Characterization of HicAB toxin-antitoxin module of Sinorhizobium meliloti. *BMC Microbiol* 19:10. doi: 10.1186/s12866-018-1382-6

Tsilibaris, V., Maenhaut-Michel, G., Mine, N., and Van Melderen, L. (2007). What is the benefit to *Escherichia coli* of having multiple toxin-antitoxin systems in its genome? *J Bacteriol* 189, 6101–6108. doi: 10.1128/JB.00527-07

Van Melderen, L. (2010). Toxin-antitoxin systems: why so many, what for? *Curr Opin Microbiol* 13, 781–785. doi: 10.1016/j.mib.2010.10.006

Van Melderen, L., and Saavedra De Bast, M. (2009). Bacterial toxin-antitoxin systems: more than selfish entities? *PLoS Genet* 5:e1000437. doi: 10.1371/journal.pgen.1000437

Vang Nielsen, S., Turnbull, K. J., Roghanian, M., Baerentsen, R., Semanjski, M., Brodersen, D. E., et al. (2019). Serine-Threonine Kinases Encoded by Split hipA Homologs Inhibit Tryptophanyl-tRNA Synthetase. *MBio* 10, e1138–e1119. doi: 10.1128/mBio.01138-19

Verma, M., Choi, J., Cottrell, K. A., Lavagnino, Z., Thomas, E. N., Pavlovic-Djuranovic, S., et al. (2019). A short translational ramp determines the efficiency of protein synthesis. *Nat Commun* 10, 5774. doi: 10.1038/s41467-019-13810-1

Verma, S., and Bhatnagar, R. (2014). MoxT toxin of Bacillus anthracis exhibits sequence specific ribonuclease activity. *Biochem Biophys Res Commun* 450, 998–1004. doi: 10.1016/j.bbrc.2014.06.092

Vesper, O., Amitai, S., Belitsky, M., Byrgazov, K., Kaberdina, A. C., Engelberg-Kulka, H., et al. (2011). Selective translation of leaderless mRNAs by specialized ribosomes generated by MazF in *Escherichia coli*. *Cell* 147, 147–157. doi: 10.1016/j.cell.2011.07.047

Vetting, M. W., Bareich, D. C., Yu, M., and Blanchard, J. S. (2008). Crystal structure of RimI from *Salmonella* typhimurium LT2, the GNAT responsible for N(alpha)-acetylation of ribosomal protein S18. *Protein Sci* 17, 1781–1790. doi: 10.1110/ps.035899.108

Veyron, S., Peyroche, G., and Cherfils, J. (2018). FIC proteins: from bacteria to humans and back again. *Pathog Dis* 76, fty012. doi: 10.1093/femspd/fty012

Wade, J. T., and Laub, M. T. (2019). Concerns about "Stress-Induced MazF-Mediated Proteins in *Escherichia coli*". *MBio* 10, e825–e819. doi: 10.1128/mBio.00825-19

Walling, L. R., and Butler, J. S. (2018). Homologous VapC Toxins Inhibit Translation and Cell Growth by Sequence-Specific Cleavage of tRNA(fMet). *J Bacteriol* 200, e582–e517. doi: 10.1128/JB.00582-17

Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., and Barton, G. J. (2009). Jalview Version 2–a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–1191. doi: 10.1093/bioinformatics/btp033

Wei, Y., Gao, Z. Q., Otsuka, Y., Naka, K., Yonesaki, T., Zhang, H., et al. (2013). Structure-function studies of *Escherichia coli* RnlA reveal a novel toxin structure involved in bacteriophage resistance. *Mol Microbiol* 90, 956–965. doi: 10.1111/mmi.12409

Wen, Y., Behiels, E., Felix, J., Elegheert, J., Vergauwen, B., Devreese, B., et al. (2014). The bacterial antitoxin HipB establishes a ternary complex with operator DNA and phosphorylated toxin HipA to regulate bacterial persistence. *Nucleic Acids Res* 42, 10134–10147. doi: 10.1093/nar/gku665

Wendrich, T. M., Blaha, G., Wilson, D. N., Marahiel, M. A., and Nierhaus, K. H. (2002). Dissection of the mechanism for the stringent factor RelA. *Mol Cell* 10, 779–788.

Whisstock, J. C., and Lesk, A. M. (1999). SH3 domains in prokaryotes. *Trends Biochem Sci* 24, 132–133. doi: 10.1016/s0968-0004(99)01366-3

Wilcox, B., Osterman, I., Serebryakova, M., Lukyanov, D., Komarova, E., Gollan, B., et al. (2018). *Escherichia coli* ItaT is a type II toxin that inhibits translation by acetylating isoleucyl-tRNAIle. *Nucleic Acids Res* 46, 7873–7885. doi: 10.1093/nar/gky560

Winther, K., Tree, J. J., Tollervey, D., and Gerdes, K. (2016). VapCs of Mycobacterium tuberculosis cleave RNAs essential for translation. *Nucleic Acids Res* 44, 9860–9871. doi: 10.1093/nar/gkw781

Winther, K. S., Brodersen, D. E., Brown, A. K., and Gerdes, K. (2013). VapC20 of Mycobacterium tuberculosis cleaves the sarcin-ricin loop of 23S rRNA. *Nat Commun* 4, 2796. doi: 10.1038/ncomms3796

Winther, K. S., and Gerdes, K. (2011). Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA. *Proc Natl Acad Sci U S A* 108, 7403–7407. doi: 10.1073/pnas.1019587108

Wozniak, R. A., and Waldor, M. K. (2009). A toxin-antitoxin system promotes the maintenance of an integrative conjugative element. *PLoS Genet* 5:e1000439. doi: 10.1371/journal.pgen.1000439

Xu, K., Dedic, E., and Brodersen, D. E. (2016). Structural analysis of the active site architecture of the VapC toxin from *Shigella* flexneri. *Proteins* 84, 892–899. doi: 10.1002/prot.25002

Yamaguchi, Y., Nariya, H., Park, J. H., and Inouye, M. (2012). Inhibition of specific gene expressions by protein-mediated mRNA interference. *Nat Commun* 3, 607. doi: 10.1038/ncomms1621

Yamaguchi, Y., Park, J. H., and Inouye, M. (2009). MqsR, a crucial regulator for quorum sensing and biofilm formation, is a GCU-specific mRNA interferase in *Escherichia coli*. *J Biol Chem* 284, 28746–28753. doi: 10.1074/jbc.M109.032904

Yamanishi, H., and Yonesaki, T. (2005). RNA cleavage linked with ribosomal action. *Genetics* 171, 419–425. doi: 10.1534/genetics.105.042515

Yao, J., Guo, Y., Zeng, Z., Liu, X., Shi, F., and Wang, X. (2015). Identification and characterization of a HEPN-MNT family type II toxin-antitoxin in Shewanella oneidensis. *Microb Biotechnol* 8, 961–973. doi: 10.1111/1751-7915.12294

Yuan, J., Sterckx, Y., Mitchenall, L. A., Maxwell, A., Loris, R., and Waldor, M. K. (2010). *Vibrio cholerae* ParE2 poisons DNA gyrase via a mechanism distinct from other gyrase inhibitors. *J Biol Chem* 285, 40397–40408. doi: 10.1074/jbc.M110.138776

Yusupov, M. M., Yusupova, G. Z., Baucom, A., Lieberman, K., Earnest, T. N., Cate, J. H., et al. (2001). Crystal structure of the ribosome at 5.5 A resolution. *Science* 292, 883–896. doi: 10.1126/science.1060089

Zhang, D., de Souza, R. F., Anantharaman, V., Iyer, L. M., and Aravind, L. (2012). Polymorphic toxin systems: Comprehensive characterization of trafficking modes, processing, mechanisms of action, immunity and ecology using comparative genomics. *Biol Direct* 7, 18. doi: 10.1186/1745-6150-7-18

Zhang, J., Zhang, Y., Zhu, L., Suzuki, M., and Inouye, M. (2004). Interference of mRNA function by sequence-specific endoribonuclease PemK. *J Biol Chem* 279, 20678–20684. doi: 10.1074/jbc.M314284200

Zhang, Y., Yamaguchi, Y., and Inouye, M. (2009). Characterization of YafO, an *Escherichia coli* toxin. *J Biol Chem* 284, 25522–25531. doi: 10.1074/jbc.M109.036624

Zhang, Y., Zhu, L., Zhang, J., and Inouye, M. (2005). Characterization of ChpBK, an mRNA interferase from *Escherichia coli*. *J Biol Chem* 280, 26080–26088. doi: 10.1074/jbc.M502050200

Zhu, L., Inoue, K., Yoshizumi, S., Kobayashi, H., Zhang, Y., Ouyang, M., et al. (2009). Staphylococcus aureus MazF specifically cleaves a pentad sequence. *J Bacteriol* 191, 3248–3255. doi: 10.1128/JB.01815-08

Zhu, L., Phadtare, S., Nariya, H., Ouyang, M., Husson, R. N., and Inouye, M. (2008). The mRNA interferases, MazF-mt3 and MazF-mt7 from Mycobacterium tuberculosis target unique pentad sequences in single-stranded RNA. *Mol Microbiol* 69, 559–569. doi: 10.1111/j.1365-2958.2008.06284.x

Zhu, L., Zhang, Y., Teh, J. S., Zhang, J., Connell, N., Rubin, H., et al. (2006). Characterization of mRNA interferases from Mycobacterium tuberculosis. *J Biol Chem* 281, 18638–18643. doi: 10.1074/jbc.M512693200

Zorzini, V., Buts, L., Sleutel, M., Garcia-Pino, A., Talavera, A., Haesaerts, S., et al. (2014). Structural and biophysical characterization of Staphylococcus aureus SaMazF shows conservation of functional dynamics. *Nucleic Acids Res* 42, 6709–6725. doi: 10.1093/nar/gku266

Zorzini, V., Mernik, A., Lah, J., Sterckx, Y. G., De Jonge, N., Garcia-Pino, A., et al. (2016). Substrate Recognition and Activity Regulation of the *Escherichia coli* mRNA Endonuclease MazF. *J Biol Chem* 291, 10950–10960. doi: 10.1074/jbc.M116.715912

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

ORIGINAL RESEARCH
published: 18 August 2020
doi: 10.3389/fgene.2020.00856

Modulation of *Escherichia coli* Translation by the Specific Inactivation of tRNA^{Gly} Under Oxidative Stress

*Lorenzo Eugenio Leiva[1], Andrea Pincheira[1], Sara Elgamal[2], Sandra D. Kienast[3,4,5], Verónica Bravo[6], Johannes Leufken[3,4,5], Daniela Gutiérrez[1], Sebastian A. Leidel[3,4,5], Michael Ibba[2] and Assaf Katz[1]**

[1]*Programa de Biología Celular y Molecular, ICBM, Facultad de Medicina, Universidad de Chile, Santiago, Chile,* [2]*Department of Microbiology and The Center for RNA Biology, The Ohio State University, Columbus, OH, United States,* [3]*Max Planck Research Group for RNA Biology, Max Planck Institute for Molecular Biomedicine, Münster, Germany,* [4]*Cells-in-Motion Cluster of Excellence and Faculty of Medicine, University of Münster, Münster, Germany,* [5]*Research Group for RNA Biochemistry, Department of Chemistry and Biochemistry, University of Bern, Bern, Switzerland,* [6]*Unidad de Microbiología, Escuela de Medicina, Facultad de Ciencias Médicas, Universidad de Santiago de Chile, Santiago, Chile*

OPEN ACCESS

Edited by:
Joseph Curran,
Université de Genève, Switzerland

Reviewed by:
Dominique Belin,
Université de Genève, Switzerland
Dieter Soll,
Yale University, United States

***Correspondence:**
Assaf Katz
askatz@uchile.cl

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *15 December 2019*
Accepted: *14 July 2020*
Published: *18 August 2020*

Citation:
Leiva LE, Pincheira A, Elgamal S, Kienast SD, Bravo V, Leufken J, Gutiérrez D, Leidel SA, Ibba M and Katz A (2020) Modulation of Escherichia coli Translation by the Specific Inactivation of tRNA^{Gly} Under Oxidative Stress. Front. Genet. 11:856. doi: 10.3389/fgene.2020.00856

Bacterial oxidative stress responses are generally controlled by transcription factors that modulate the synthesis of RNAs with the aid of some sRNAs that control the stability, and in some cases the translation, of specific mRNAs. Here, we report that oxidative stress additionally leads to inactivation of tRNA^{Gly} in *Escherichia coli*, inducing a series of physiological changes. The observed inactivation of tRNA^{Gly} correlated with altered efficiency of translation of Gly codons, suggesting a possible mechanism of translational control of gene expression under oxidative stress. Changes in translation also depended on the availability of glycine, revealing a mechanism whereby bacteria modulate the response to oxidative stress according to the prevailing metabolic state of the cells.

Keywords: tRNA^{Gly}, *Escherichia coli*, oxidative stress, codon, glycine, translation

INTRODUCTION

Bacteria, like other organisms, need to adapt to environmental conditions that are constantly changing. Some of these conditions induce oxidative stress in bacteria due to either an increase in oxidants or a decrease in the ability of bacteria to defend against them. In the absence of an adequate protective response to oxidative stress, numerous macromolecules may be damaged, including proteins, lipids, and nucleic acids (Imlay, 2008).

The response to oxidative stress has been extensively studied, in particular because generation of an oxidative attack by macrophages and polymorphonuclear leukocytes is one of the main defense strategies of the human body against invading bacteria once they have crossed the primary physical barriers (Slauch, 2011; Nguyen et al., 2017). In order to adapt to conditions that induce oxidative stress, bacteria may: (I) reduce motility, increase exopolysacharide production, and induce biofilm formation, thereby reducing accessibility to molecules that produce oxidative stress (Gambino and Cappitelli, 2016); (II) inhibit replication, preventing DNA mutagenesis at sites of base oxidation (Imlay, 2013) and possibly

also reducing the toxicity of replication near the site of repair for oxidized bases (Charbon et al., 2014); (III) reduce the rate of global translation (Katz and Orellana, 2012; Zhong et al., 2015; Zhu and Dai, 2019); (IV) selectively induce the production of several proteins involved in the reduction, repair, or degradation of oxidant molecules or oxidized biological targets such as thiol groups and Fe/S clusters in proteins (Imlay, 2013); and (V) decrease the production of reduced nicotinamide adenine dinucleotide (NADH) in order to increase production of reduced nicotinamide adenine dinucleotide phosphate (NADPH), required to reduce oxidant molecules and oxidized targets (Rui et al., 2010; Shen et al., 2013). Finally, oxidative stress also induces some members from a bacterial community to enter a partially quiescent state known as "persistence" (Wu et al., 2012), where several primary metabolic pathways are repressed and stress responses are induced (Lewis, 2010; Cohen et al., 2013).

Transcriptional control of gene expression plays a major role in coordinating the cellular response to oxidative stress. In *Escherichia coli* and other enterobacteria, OxyR and SoxR sense the presence of hydrogen peroxide and oxygen superoxide, respectively, regulating the transcription of genes that belong to each regulon (Imlay, 2013). These transcription factors are aided by others, like Fur or Fnr, that modulate more specific aspects of the response (Chiang and Schellhorn, 2012; Shimizu, 2016). Beyond this "transcription focused" response, reports from diverse laboratories suggest that the defense mechanisms against oxidative stress also depend on translation, for example, *via* regulation of transfer RNA (tRNA) metabolism and translation elongation. It has been shown that changes in error rate of aminoacyl-tRNA synthetases are altered by oxidation of editing domains that may increase or decrease their activities under oxidative stress (Ling and Söll, 2010; Wu et al., 2014; Steiner et al., 2019). Others have found that tRNA may be oxidized (Liu et al., 2012) and, in at least some *E. coli* strains, oxidative stress induces a general and undiscriminated degradation of tRNAs that strongly reduce translation elongation, eventually provoking cell death (Zhong et al., 2015; Zhu and Dai, 2019). Others have found that, in *E. coli* and other bacteria, oxidation of translation elongation factors may also inhibit elongation (Kojima et al., 2009; Nagano et al., 2015; Yutthanasirikul et al., 2016).

To investigate whether changes in the concentration of particular tRNAs may modulate the translation of the *E. coli* transcriptome, we screened for changes in the levels of functional tRNAs under sub-lethal oxidative stress. Unexpectedly and in contrast to other reports, we observed that in the strain used in this research, *E. coli* K-12 MG1655, only tRNA^{Gly} is inactivated under oxidative stress while nine other tRNAs remain active and at unchanged concentrations. Over production of specific tRNA^{Gly} isoacceptors altered the response of *E. coli* to oxidative stress, altering motility, carbohydrate consumption, and growth kinetics. The observed changes directly correlated with alterations in the translation efficiency of Gly codons under oxidative stress, suggesting a mechanism by which changes in active tRNA^{Gly} levels may regulate the response to oxidative stress.

MATERIALS AND METHODS

Strains and Culture Media

All experiments performed in this work used wild-type (WT) *E. coli* K-12 MG1655 strain. Strains were cultured in either lysogeny broth (LB) media (1% tryptone, 0.5% yeast extract, and 0.5% NaCl), M9 media (47.7 mM Na_2HPO_4, 22.0 mM KH_2PO_4, 8.6 mM NaCl, 18.7 mM NH_4Cl, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, and 0.4% Glycerol), or low phosphate media (40 mM MOPS, 4 mM Tricine, 50 mM KCl, 10 mM NH_4Cl, 0.2 mM KH_2PO_4, 0.5 mM $MgSO_4$, 10 μM $FeCl_3$, and 0.4% Glucose) supplemented with branched amino acids. When indicated, isopropyl β-d-1-thiogalactopyranoside (IPTG; 100 μM), branched amino acids (isoleucine, leucine, and valine 50 μg/ml each), Gly (50 μg/ml glycine), diverse sugars (glucose, arabinose, lactose, or manose, 0.2–0.4%), phenol red (0.1 μg/ml), ampicillin (100 μg/ml), or paraquat (up to 1M) were added to the culture media.

tRNA Purification

tRNA was extracted using previously published protocols (Raczniak et al., 2001) with increased 2-mercaptoethanol added to quench possible remnants of oxidizing molecules. 250 ml of LB were inoculated with 1.5 ml of an overnight culture of *E. coli*, incubated at 37°C, and shaken at 225 rpm. At an OD_{600} of 0.6–0.7, H_2O_2 or paraquat was added for a final concentration of 2.5 and 1 mM, respectively. After 20 min (H_2O_2) or 30 min (paraquat) of incubation, the culture was pelleted at 11,000 *g* for 5 min at 37°C. Cells were resuspended in 2.5 ml of buffer A (20 mM Tris HCl, pH 7.0; 20 mM $MgCl_2$; and 20 mM 2-mercaptoethanol) and extracted by shaking for 20 min at room temperature with 2.5 ml of acid phenol. Aqueous phase was recovered after centrifugation at 6,500 *g* for 10 min at room temperature and stored at 4°C. Phenol was re-extracted with additional 2.5 ml of buffer A. Both buffer A extracts were mixed together and re-extracted with 5 ml of acid phenol. The aqueous phase was recovered. Isopropanol was added to a final concentration of 20% and was centrifuged at 9,000 *g* for 60 min. The supernatant was recovered and isopropanol concentration was adjusted to 60%. The mixture was centrifuged at 11,000 *g* for 60 min. Supernatant was discarded, the pellet briefly dried and then dissolved in 1.25 ml of 200 mM Tris acetate pH 8.5 plus 20 mM 2-mercaptoethanol and incubated for 60 min at 37°C to deacylate tRNAs. Samples were further purified in a DE52 or DEAE sepharose column (~250 μl resin). Sample was loaded in the column and subsequently cleaned with 50 volume buffer II (20 mM Tris HCl pH 7.0; 200 mM NaCl and 5 mM 2-mercaptoethanol). tRNA was eluted with buffer III (20 mM Tris HCl, pH 7.0; 1M NaCl; and 5 mM 2-mercaptoethanol). tRNA from fractions with higher absorbance at 260 nm was recovered by precipitating with 0.1 volume 3 M sodium acetate pH 4.5 and 2 volume ethanol. Samples was stored at −20°C for at least 30 min and then centrifuged at 9,000 *g* for 1 h. Pellets were cleaned with cold 80% ethanol and resuspended in H_2O.

tRNA samples used for mass spectrometry analyses were further purified using biotinylated beads. For these samples, 180 μg of total tRNA were dissolved in hybridization solution

[0.1× saline sodium citrate (SSC) buffer, 0.1% sodium dodecyl sulfate (SDS), and 5 mM 2-mercaptoethanol] containing 0.5 mM EDTA and 2.5 μM of the corresponding biotinylated probe. We used the same biotinylated probes as for Northern blots (**Supplementary Table S3**). The mixture was incubated for 2 min at 90°C and then rapidly cooled to 41°C. Samples were further incubated for 120 min at this temperature. Samples were then mixed with 60 μl streptavidin/sepharose beads (in 150 μl hybridization buffer) and incubated for 30 min at this temperature while shaking at 1 min intervals. Samples were centrifuged for 20 s at 3,000 *g* and the supernatant was eliminated. Then samples were cleaned eight times with hybridization buffer (3 min incubations at 41°C with gentle shaking each min. Spin 20 s at 3,000 *g* to eliminate supernatant) and eluted in 40 μl of the same buffer at 80°C.

tRNA Quantification by Aminoacylation

tRNA's concentration was estimated from the plateau of an aminoacylation reaction progress curve at 37°C. Reactions were started by adding tRNA extracts to get 0.2 μg/μl in a mixture containing 1× reaction buffer (100 mM HEPES KOH, pH 7.2; 30 mM KCl; and 12 mM $MgCl_2$), 5 mM ATP pH 7.0, 10 mM 2-mercaptoethanol, 8 U/ml pyrophosphatase (Roche 10 108 987 001), 2.7 mg/ml of a S100 extract from *E. coli* K-12 MG1655 (cleaned using DE52 resin to eliminate amino acids and RNA), and a mix of non-radioactive and ^{14}C amino acid (**Supplementary Table S4** for final concentrations). At defined time points, 7 μl of aliquots were deposited in filter paper saturated in 5% trichloroacetic acid to precipitate aminoacyl-tRNAs. Papers were washed at room temperature three times for 5 min in 5% trichloroacetic acid and once in 100% ethanol. Then, papers were dried and aminoacyl-tRNAs were quantified in a scintillation counter. Background was subtracted based on experiments where no tRNA was added to account for non-specific binding of radioactive amino acids to filter papers and the potential tRNA traces present from S100 extracts.

RNA Mass Spectrometry

The detailed protocol for the analysis of RNA by mass spectrometry has been described elsewhere (Sarin et al., 2018). In brief, tRNAs were digested to single nucleosides essentially, as previously described (Alings et al., 2015). Chromatographic separations of the samples were performed using a self-packed Hypercarb capillary column (75 μm ID × 500 mm) coupled to a Proxeon EASY nLC (Thermo Fisher Scientific GmbH, Dreieich, Germany). Samples were separated using solvent A (5 mM ammonium formate pH 5.2) and solvent B (acetonitrile) in a multi-step gradient (2–20% B for 3 min; 20–75% B for 3–50 min; 75–100% B for 5 min; hold at 100% B for 15 min). Subsequently, samples were analyzed using a Q Exactive Mass Spectrometer (Thermo Finnigan LLC, San Jose, CA).

Quantitative Analysis of LC-MS/MS Data

Thermo RAW files were converted to the mzML format (Martens et al., 2011) using msConvert as part of ProteoWizard (version 3.0.10738; Chambers et al., 2012). Quantitative data analysis was performed using pymzML (version 2.0.0; Bald et al., 2012) and pyQms (version 0.5.0; Leufken et al., 2017). Chemical formulae of all nucleosides (including modified forms) were retrieved from the MODOMICS database (Boccaletto et al., 2017).

pyQms was used to calculate high-accuracy isotopologue patterns for all chemical formulas, and these patterns were matched onto all MS1 spectra. Quantification of nucleosides for individual samples is based on the maximum intensity of the matched isotope pattern chromatogram (MIC). To assess quantification quality, pyQms calculates a weighted similar match score (mScore; Leufken et al., 2017). Detection and quantitation of selected nucleosides were manually validated.

Total RNA Extraction

2 ml of bacterial cultures in LB or 5 ml from cultures in M9 media were pelleted for 1 min at 12,000 *g*. Pellet were resuspended in 50 μl lysis buffer (83 mM Tris HCl, pH 6.8, 18 mM EDTA pH 8, 1.7% SDS, and 1.6% 2-mercaptoethanol) and incubated for 3 min at 37°C. 1.5 ml of TRIzol was added, and total RNA was extracted following the manufacturer's protocol.

Northern Blot Assay

Most Northern blot analyses were performed using biotinylated probes. For some RNA, we additionally used a non-labeled probe to help "unwind" the tRNA structures (list of probes in **Supplementary Table S3**). Samples were transferred to positively charged nylon membranes for 2 h at 22 volts in 0.5× TBE. Then, RNA was fixated by UV radiation and membranes were blocked for 30 min at the temperature indicated in **Supplementary Table S3** in hybridization solution (6× SSC, 70 μg/ml heat-denatured salmon sperm DNA, 0.1% SDS, and 5× Denhardt's solution). After blocking, probes were added directly to the same solution and incubated overnight at the same temperature. Membrane was then washed for 3 min at room temperature with solution A (2× SSC and 0.1% SDS) and then twice for 15 min at the temperature specified in **Supplementary Table S3** in solution B (0.1× SSC and 0.1% SDS). After this treatment, the membrane was blocked for 30 min at room temperature with a blocking solution [1% casein in maleic buffer (0.1M maleic acid and 0.15M NaCl pH 7.5)]. Then, 0.1 μg/ml of streptavidin-horseradish peroxidase was added to the blocking solution. Membranes were incubated for 30 min at room temperature and then washed twice for 15 min with maleic acid buffer plus 0.3% (V/V) tween-20. Finally, the membranes were washed for 3 min in pre-detection buffer (0.1M Tris HCl, 0.1M NaCl, pH 9.5) and developed using a chemiluminescent kit (SuperSignal West Pico Chemiluminescent Substrate, Prod#34080). Determination of aminoacylation levels in strains overproducing tRNAs was performed using ^{32}P labeled probes. A similar protocol was used, but images were acquired using phosphorimager technology.

Determination of the *in vivo* Levels of tRNA Aminoacylation

Total RNA was purified in acidic conditions, and then, the 3' extreme nucleotide was eliminated by sodium periodate

oxidation followed by β-elimination following previously described protocols (Salazar et al., 2004). Briefly, 50 ml of LB were inoculated with 300 µl on preinoculum of *E. coli* and incubated at 37°C and constant shaking. At an OD_{600} of 0.9–1.0, 15 ml of culture was pelleted at 10,000 *g* for 6 min. Paraquat to 1 mM was added to the remaining culture and continued incubating for 30 min, after which 15 ml were similarly pelleted. Immediately after pelleting, each bacterial pellet was resuspended in 500 µl of 0.3 M sodium acetate pH 5.2 with 1 mM EDTA pH 8.0. After resuspension, 500 µl of acid phenol were added and the mixture was incubated for 10 min on ice with intermittent mixing. Then, phases were separated by centrifuging for 6 min at 10,000 *g*. Aqueous phase was recovered and RNA was precipitated by adding 1 ml of ethanol and storing samples at −80°C. After samples from stressed cells were kept for 30 min at −80°C, all samples were centrifuged (14,000 *g*, 30 min). Pellets were washed with 0.5 ml of 75% ethanol with 10 mM sodium acetate pH 5.2 and then resuspended in 50 µl of H_2O. It is recommended not to use freshly distilled water to allow pH of water to decrease by absorption of atmospheric CO_2. Each sample was divided into two 25 µl aliquots. 1.42 µl of 3 M sodium acetate of pH 5.2 was added to tubes "A" that were then stored at −80°C. tRNA in tubes "B" was deaminoacylated by adding 6.25 µl of 1 M Tris acetate pH 9.0 and incubating for 60 min at 37°C. Samples in tubes B were precipitated by adding 3.13 µl of 3 M sodium acetate of pH 5.2 and 62.5 µl ethanol. Samples were stored at −80°C for at least 30 min, after which samples were centrifuged (30 min, 13,000 rpm). Pellets were washed with ethanol 70%, dried, and resuspended in 26.4 µl of 160 mM sodium acetate pH 5.2. Samples A were thawed and 4.84 µl of freshly prepared 250 mM sodium periodate was added to tubes A and B. Tubes were wrapped in aluminum foil and incubated for 90 min on ice. Then, 12.97 µl of 20% glucose was added. After an additional 90 min incubation in ice, 4.3 µl of 3 M sodium acetate pH 5.2 and 87 µl ethanol were added. Samples were stored at least for 30 min at −80°C and centrifuged (30 min, 13,000 rpm). Pellets were resuspended in 250 µl of 0.5M lysine pH 8.0 and incubated for 60 min at 45°C. Then, 25 µl of 3M sodium acetate pH 5.2 and 500 µl ethanol were added, and samples were stored at least for 30 min at −80°C. Tubes were then centrifuged (30 min, 13,000 rpm), and after washing with 70% ethanol, pellets were dried and resuspended in 15 µl water. Then, samples were analyzed by 10% polyacrylamide gels with 8 M urea and by Northern blot analysis.

Cloning and Mutation Protocols

Plasmid pBAD30SFIT (Rojas et al., 2018) contains a tandem fluorescent transcriptional fusion cassette composed of superfold green fluorescent protein (sfGFP) followed directly by a modified mCherry, itag-mCherry (Henriques et al., 2013). The plasmid contains a XhoI-SpeI site after the third codon of *sfgfp*, where tetra codon sequences were inserted using annealed oligo cloning with the oligonucleotide pairs described in **Supplementary Table S5**. *narJ* was also cloned in XhoI-SpeI restriction sites after amplifying the gene from *E. coli* K-12 MG1655 genomic DNA using external oligonucleotides NarJ_EcoRI_5'_Fw and NarJ_XhoI_3'_GGA_Rv (**Supplementary Table S6**). To clone the mutant version of *narJ*, a similar protocol was used, exchanging primer NarJ_XhoI_3'_GGA_Rv by NarJ_XhoI_3'_GGC_Rv. The same protocol was used to clone these genes in pBAD30SFIT-HP, which codes for a hairpin between *gfp* and *mCherry*. The hairpin was introduced in pBAD30SFIT by annealed oligo cloning of oligonucleotides Str_Yam_Fw and Str_Yam_Rv (**Supplementary Table S6**) in BglII y PciI restriction sites.

Cloning of tRNA genes was performed by annealed oligo cloning of oligonucleotides listed in **Supplementary Table S7** between EcoRI and HindIII sites of pKK223-3.

E. coli K-12 MG1655 Δ*glyVX*::FRT was constructed by the Red-swap method (Datsenko and Wanner, 2000) using primers glyV (H1 + P1) and glyX (H2 + P2; **Supplementary Table S8**), as well as plasmid pLZ01 (Blondel et al., 2013) as template for amplification of a Cam resistance cassette flanked by the FRT sites (FLP recombinase target sequence).

Translation Efficiency Analyses

M9 media supplemented with branched amino acids (50 µg/ml each) and ampicillin (100 µg/ml) were inoculated with bacteria from a saturated overnight culture in similar media and grown at 37°C in an orbital shaker. When indicated, the inoculated media were also supplemented with Gly (50 µg/ml). When bacteria reached mid-log phase (OD_{600} ~0.4–0.6), a 50 µl aliquot of it was diluted in a 96-well optical-bottom plate with 150 µl fresh M9 media supplemented with arabinose (0.4% final concentration). When indicated, media additionally contained paraquat (700 µM final concentration). Plates were further shaken for 2 h at 37°C. Then, OD_{600} and fluorescence intensity of green fluorescent protein (GFP; Ex. 480 ± 4.5 nm, Em. 515 ± 10 nm) and mCherry (Ex. 555 ± 4.5 nm, Em. 600 ± 10 nm) were measured in a microplate reader (Infinite M200 PRO, Tecan). When comparing GFP/mCherry fluorescence ratios of different strains (WT vs. Δ*glyVX*::FRT), data were normalized by the GFP/mCherry fluorescence ratio of a control without additional codons (S1).

Motility Assay

Strains were cultured in LB media supplemented with 100 µg/ml ampicillin and 100 µM IPTG. When the culture reached an OD_{600} of ~0.7, an aliquot was centrifuged and the pellet was resuspended in LB to an OD_{600} of 1. 5 µl of these samples were used to inoculate LB plates with 0.3% agar, 100 µg/ml ampicillin and 100 µM IPTG. When indicated, 500 µM paraquat were also added to the plates. The plates were incubated during 8 (control) or 24 h (paraquat) in a humid chamber at 30°C. After this time, the radial growth was measured (Ha et al., 2014) and expressed as a ratio to colony diameter of the strain carrying the empty plasmid (62 ± 5 mm under control conditions and 19 ± 10 mm when 500 µM paraquat was added). These experiments could not be performed in plates with M9 media because motility was too low for all strains when paraquat was added.

Growth Curves

E. coli K-12 MG1655 was transformed with plasmids carrying the genes for each tRNA^{Gly} isoacceptor and stored at −80°C. A fresh ON culture of these strains in LB media supplemented with 100 μg/ml ampicillin and 100 μM IPTG was diluted to DO_{600nm} ~0.05 in a similar media with or without 1 mM paraquat and transferred to a 96-well microtiter plate. Growth was subsequently followed in a thermostated microplate reader (Infinite M200 PRO, Tecan) at 37°C and intercalating orbital (142 rpm, 6 mm amplitude) and linear (296 rpm, 6 mm amplitude) shaking in 10 min intervals. Using strains that have been cultured several times in LB plates gave inconsistent results, suggesting that strains overproducing tRNA^{Gly} accumulate mutations that altered their behavior.

Carbohydrate Fermentation Assay

ON cultures of bacteria grown in M9 media supplemented with glycerol 0.4%, Gly 50 μg/ml, branched amino acid of 50 μg/ml, and ampicillin 100 μg/μl were used to inoculate similar media and incubated at 37°C. When cultures reached an OD_{600} value of 0.6–0.8, a sample was diluted to an OD_{600} of 0.2 in similar media. Around 50 μl of these samples were mixed in a 96-well plate with 50 μl of similar media plus the corresponding carbohydrate (0.4% final concentration), IPTG (100 μM final concentration), and phenol red (0.1 μg/L) as pH indicator. Media also had diverse concentrations of paraquat (0–350 μM). Samples were incubated for 12 h at 37°C, after which plates were centrifuged (3,000 rpm, 7 min). Absorbance of the supernatant was measured at 560 nm.

Paraquat MIC Determination

MIC values were estimated using previously described methods (Wiegand et al., 2008). Briefly, 50 μl of culture with 10^8 CFU in M9 media supplemented with glycerol 0.4%, Gly 50 μg/ml, branched amino acids of 50 μg/ml each, 100 μM IPTG, and 100 μg/ml ampicillin were mixed with 50 μl of the same media containing serial dilutions of paraquat in 96-well plates. Plates were incubated in ON at 37°C, and then bacterial growth was determined by OD_{600}.

Analysis of Codon Usage

Codon usage for each gene was calculated with an in-house Perl 5 script and using the current RefSeq annotation for *E. coli* K-12 MG1655 genes (RefSeq assembly accession: GCF_000005845.2). Enrichment analysis was performed in EcoCyc platform (Keseler et al., 2017).

RESULTS

Oxidative Stress Induces a Decrease in the Levels of Active tRNA^{Gly}

Aminoacylated tRNAs (aa-tRNAs) are essential for elongation of the nascent peptide during mRNA translation. While translation is mainly regulated at the initiation step, changes in the modification status and/or aminoacylation levels of tRNA can regulate translation by altering elongation rates (Starzyk, 1984; Subramaniam et al., 2013a, 2014; Katz et al., 2016; Zhu and Dai, 2019). To determine the role of tRNA alterations in the bacterial oxidative stress response, we screened for changes in concentrations of active tRNAs upon exposure to oxidants using the tRNA aminoacylation reaction. Total tRNA was purified from *E. coli* K-12 MG1655 (Blattner et al., 1997) cells cultivated under control conditions or oxidative stress induced by addition of 1 mM paraquat or 2.5 mM H_2O_2. Addition of paraquat led to a minor decrease in growth, while exposure to H_2O_2 arrested cell growth for ~2 h, after which cells resumed replication (**Supplementary Figure S1**). Purified total tRNA was deaminoacylated and subsequently selectively aminoacylated with 10 available radioactive amino acids using cell-free extracts from *E. coli* cultured in control conditions. This allowed the screening of changes in the levels of tRNAs responsible for the decoding of 10 different amino acids during translation. Out of 10 tested tRNAs, only tRNA^{Gly} showed a statistically significant decrease in the levels of active tRNA after stress by exposure to either paraquat or H_2O_2 (**Figure 1A**). Previous reports have shown a general decrease in total tRNA levels under oxidative stress in minimal media (Zhong et al., 2015; Zhu and Dai, 2019). In contrast, we observed that tRNA levels remained fairly constant (**Figure 1B**), suggesting the observed decrease in the levels of active tRNA^{Gly} was a specific response to oxidative stress. Since we observe comparable effects on tRNA^{Gly} after addition of either H_2O_2 or paraquat, we confined additional studies to the effects of paraquat, that is continually reduced by the cellular NADPH pool (Hassan and Fridovich, 1979) producing a constant, readily reproducible oxidative stress. Furthermore, as oxidative stress is known to inactivate dihydroxyacid dehydratase and consequentially impair the synthesis of branched amino acids (Imlay, 2008), we added Leu, Val, and Ile to cultures when using minimal media.

The Effect of Oxidative Stress on Specific tRNA^{Gly} Isoacceptors

E. coli has six-genes coding for tRNA^{Gly}. Four of them (*glyV*, *glyW*, *glyX*, and *glyY*) code for identical tRNAs that are indistinguishable by Northern blot. These tRNAs have a GCC anticodon (tRNA$^{Gly}_{GCC}$) that decodes GGC and GGU, two codons that are used at high frequency in *E. coli*. A fifth gene, *glyU*, codes for tRNA$^{Gly}_{CCC}$ that exclusively decodes GGG codons. A sixth gene, *glyT*, codes for tRNA$^{Gly}_{UCC}$ that decodes GGA and GGG codons. Codons decoded by tRNA$^{Gly}_{CCC}$ and tRNA$^{Gly}_{UCC}$ are used with a lower frequency (**Table 1** and **Supplementary Figure S2**). Genes *glyV*, *glyX*, and *glyY* are clustered together in a putative operon, while the other genes for tRNA^{Gly} are either not clustered (*glyU*) or clustered with genes for other tRNAs (*glyT* and *glyW*; **Supplementary Figure S3**). We tested whether changes in the expression levels of specific tRNA^{Gly} isoacceptors correlate with the observed differences in active tRNA levels during oxidative stress. Northern blot analyses indicated that levels of all tRNA^{Gly} isoacceptors were unaltered under oxidative stress (**Figure 1C**). The lack of variation in tRNA^{Gly} isoacceptor expression suggested that the changes in

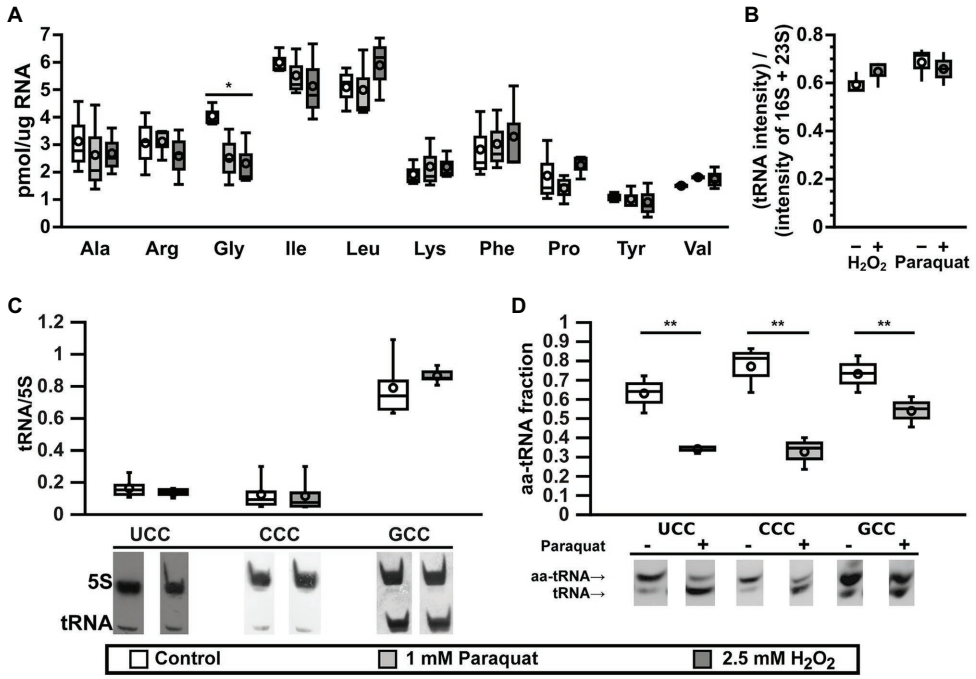

FIGURE 1 | tRNA^{Gly} are inactivated under oxidative stress. **(A)** Effect of oxidative stress on the levels of active tRNAs specific for 10 different amino acids as measured by the plateau of aminoacylation reaction. Control: lysogeny broth (LB), white bars; oxidative stress: LB + 1 mM paraquat, light gray bars; or 2.5 mM H_2O_2, dark gray bars. *$p \leq 0.01$, one way ANOVA with Dunnett versus Control posterior test for each amino acid ($n = 3$). **(B)** Effect of oxidative stress on the levels of total tRNA from cells cultured in M9 media under control (white bars) or oxidative stress (gray bars, left: 2.5 mM H_2O_2 or right: 1 mM paraquat) and analyzed by electrophoresis ($n = 3$). Intensities from tRNA are higher than expected, as ribosomal RNA (rRNA) should represent at least 80% of total RNA. We suggest this is an artifact of staining efficiency and should not alter the conclusions of the figure. **(C)** Effect of paraquat on the levels of tRNA as quantified by Northern blot of total RNA samples purified from control (LB, white bars) or stressed (LB + 1 mM paraquat, gray bars) *E. coli* cells. Data in the graph showed no significant differences using one way ANOVA ($n = 5$). **(D)** *In vivo* levels of aminoacylation of tRNA^{Gly} in *E. coli* cell before (white bars) or 30 min after stress by 1 mM paraquat (gray bars). 3' terminal nucleotide of RNAs was eliminated by oxidation with sodium periodate followed by β-elimination and analyzed by Northern blot. **$p \leq 0.05$, paired *t*-test ($n = 3$). In all box graphs, top, middle, and bottom lines of the box represent 25, 50, and 75% of the population. Whiskers represent the maximum and minimum values and the mean is represented by a circle.

TABLE 1 | tRNA^{Gly} coded in *E. coli* K-12 MG1655 genome.

tRNA^{Gly} genes	Anticodon	Decoded codon(s) (usage frequency*)	% of total tRNA**
glyV, *glyW*, *glyX*, *glyY*	GCC	GGC (2.96%) GGU (2.47%)	6.76%
glyU	CCC	GGG (1.11%)	3.31%***
glyT	UCC	GGA (0.79%) GGG (1.11%)	
-	ACC	-	

**Based on data available at GtRNAdb (Chan and Lowe, 2009).*
***As estimated in exponentially growing cells (Dong et al., 1996).*
****Spots of tRNA$^{Gly}_{CCC}$ and tRNA$^{Gly}_{UCC}$ were not resolved, so the authors reported the total abundance for both tRNA isoacceptor together.*

active tRNA^{Gly} levels might instead result from chemical modifications that impair their interaction with GlyRS during aminoacylation. We determined *in vivo* aminoacylation levels for tRNA^{Gly} directly by subjecting total tRNA pools to periodate oxidation and β-elimination followed by Northern blot (Salazar et al., 2004). Through this treatment, deaminoacylated tRNAs loses their terminal adenine and migrate faster in polyacrylamide electrophoresis gels while aa-tRNAs are protected by the amino acid and retain their original length and electrophoretic mobility. We observed decreases in the *in vivo* aminoacylation levels of the three tRNA^{Gly} isoacceptors under oxidative stress. While ~60–80% of each tRNA^{Gly} isoacceptor is aminoacylated under control conditions, after paraquat addition these levels decreased to around ~30–50% (**Figure 1D**). However, a stronger decrease was observed for tRNA$^{Gly}_{UCC}$ (from ~60 to ~30%) and tRNA$^{Gly}_{CCC}$ (from ~80 to ~30%) that are less abundant than tRNA$^{Gly}_{GCC}$ (from ~70 to ~50%). In similar experiments, we found that oxidative stress induced by paraquat produces only minor alterations to the concentration and aminoacylation levels of other tRNAs such as tRNA^{Trp}, tRNA^{Try}, and initiator tRNA^{fMet}, confirming that the observed deaminoacylation is a specific behavior of tRNA^{Gly} (**Supplementary Figure S4**). Next, we hypothesized that tRNA^{Gly} inactivation might be mediated through differences in the levels of chemical modification of the tRNA^{Gly} isoacceptors. Individual tRNA^{Gly} isoacceptors were purified and analyzed by quantitative RNA mass spectrometry. Surprisingly, while we detected all modified nucleosides that are known for each isoacceptor, we did not observe quantitative differences between the stress and the control samples (data not shown).

However, these results do not necessarily invalidate our hypothesis, as abasic sites and unknown oxidation products of modified tRNA^{Gly} nucleotides that might inactivate the tRNAs could have escaped our analysis. Finally, modifications might have been lost as a result of heating in the presence of a reducer during purification of individual tRNA^{Gly} isoacceptors.

Effect of Oxidative Stress on Translation of Gly Codons

Changes in aa-tRNA concentration are likely to alter the translation efficiency of the codons they decode. As each tRNA^{Gly} isoacceptor decodes a different set of codons, this could potentially lead to codon-dependent changes in gene expression, which in turn could give rise to distinct phenotypes. To determine whether the observed decrease in active tRNA^{Gly} affects the decoding efficiency of the corresponding Gly codons, we generated reporter constructs where we fused GFP to four contiguous identical Gly codons using mCherry as an internal control of transcript levels and global alterations of translation initiation. We used these reporters to test whether GFP production was affected in response to oxidative stress conditions where we observed changes in active tRNA^{Gly} levels and aminoacylation (**Figure 1**). We previously used this strategy to analyze the role of translation elongation factor P during the translation of several codon patterns (Elgamal et al., 2014) and the effect of natural changes in tRNA gene copy numbers on codon translation (Rojas et al., 2018). We first tested GFP production in M9 minimal medium containing glycerol and branched amino acids, both in the presence and absence of Gly. In both conditions, we observed a higher GFP/mCherry ratio when using reporters containing the frequent GGC or GGU codons as compared to the infrequently used codons (GGA or GGG). This shows that the method is sufficiently sensitive to differentiate between the translation efficiency of different Gly codons [**Figure 2**; ratio between GFP/mCherry values for the most frequently used codon (GGC) and least frequently used codon (GGA) is ~1.75 fold in absence of Gly]. Additionally, this suggests that translation of four contiguous Gly codons has a similar or slower speed than initiation that is usually considered the limiting step of translation (Subramaniam et al., 2013b; Hersch et al., 2014). Otherwise, translation of the four reporters would have produced similar amounts of GFP. We then repeated the experiment in a strain where two out of four of the genes coding for tRNA$^{Gly}_{GCC}$ (*glyVX*) were deleted, leading to a decrease in the levels of the tRNA$^{Gly}_{GCC}$ (data not shown). Decreasing the levels of tRNA$^{Gly}_{GCC}$ induced a lower translation of the *gfp* genes enriched in GGT or GGC codons (that are directly translated by tRNA$^{Gly}_{GCC}$) and only minor effects on *gfp* enriched for the two other Gly codons, indicating that the method is sensitive to changes in tRNA levels (**Supplementary Figure S5**).

While adding paraquat to minimal media at the time of reporter induction completely stopped cell replication, GFP and mCherry synthesis continued (although at much lower levels) indicating that cells can still transcribe and translate their genes following initiation of oxidative stress. The addition of paraquat induced a strong decrease in the production of both GFP and mCherry in the presence or absence of Gly (around 10 and 15 fold decrease for GFP/OD_{600} and mCherry/OD_{600} values, respectively, for the strain carrying the control plasmid; data not shown). When oxidative stress was induced in the presence of Gly, no difference was observed between the translation of each Gly codon (**Figure 2**). This result suggests that the decrease in the rate of translation initiation or another limiting step is much stronger than any effect on Gly codons translation, making differences in Gly translation unmeasurable. Instead, when Gly was absent from the culture media, differences between translation of each Gly codon were readily measurable and we observed that, unexpectedly, translation of GGA shifted from being the slower Gly codon to being the fastest codon (**Figure 2**). This suggests that under this condition, not only a fraction of tRNA^{Gly} isotypes are inactivated, but additionally, Gly becomes limiting. Thus, a reduced aminoacylation derived from low Gly availability (Böck and Neidhardt, 1966; Folk and Berg, 1970a,b; Subramaniam et al., 2013b) plus tRNA^{Gly} inactivation made Gly codon translation slow enough to produce measurable differences in GFP production. In agreement with this interpretation, if we reduce the concentration of paraquat added to the media lacking Gly from 700 to 250 μM, presumably decreasing the degree of tRNA^{Gly} inactivation, the differences between Gly codons are also not observed (**Figure 2**). Although less likely, an alternative

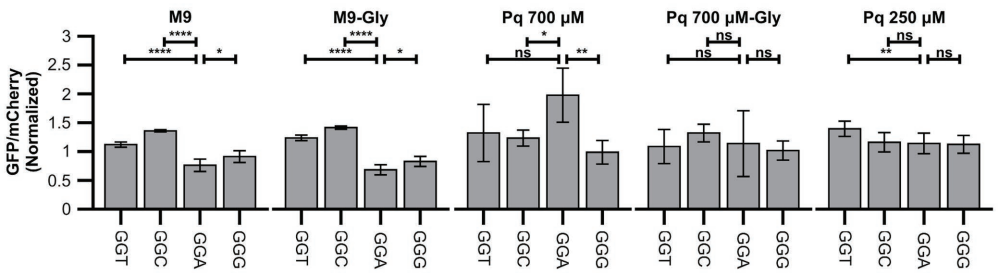

FIGURE 2 | Oxidative stress alters translation of Gly codons. Figure shows green fluorescent protein (GFP) fluorescence normalized by fluorescence of mCherry in diverse strains and conditions. Data were additionally normalized dividing by the GFP/mCherry ratio of the control strain (reporter S1, without additional codons). Four identical Gly codons were cloned in fusion to GFP and fluorescence measured in control media (M9), control media with Gly (M9-Gly), media with high concentration of paraquat in the absence (PQ 700 μM) or presence (PQ 700 μM-Gly), or media with low concentration of paraquat (PQ 250 μM). $^{****}p \leq 0.0001$, $^{**}p \leq 0.01$, $^{*}p < 0.05$, one way ANOVA with Dunnett versus GGA strain for each condition ($n \geq 3$).

interpretation that we cannot rule out is that addition of Gly somehow protects tRNA^{Gly} from inactivation without producing a similar degree of protection to translation initiation. Following this interpretation, under reduced paraquat concentrations, the fraction of inactivated tRNA^{Gly} would be much smaller than inhibition of translation initiation.

Effect of Oxidative Stress on Translation of GGA Codons in Natural Context

The results shown here indicate that oxidative stress caused by 700 μM paraquat alters translation of Gly codons when located in the context of four continued identical Gly codons within the *gfp* gene. These alterations are stronger for the most frequently used codons, transforming the most infrequently used GGA codon into the fastest Gly codon under strong oxidative stress. Nevertheless, in *E. coli* K-12 MG1655, GGA is never found in the context of four consecutive identical codons, questioning the relevance of our observation in natural genes. To verify the effect of paraquat on translation of Gly codons located in their natural context, we cloned WT *narJ* (coding for nitrate reductase molybdenum cofactor assembly chaperon) and a version where its two contiguous GGA codons were mutated to GGC in a translational fusion to *gfp* to form the *narJ(GGA)-gfp* and *narJ(GGC)-gfp* genes. Like in the previous experiments, we used mCherry as an internal control. We did not observe a significant difference between changes in GFP/mCherry fluorescence ratios under control and stress conditions for the *narJ(GGA)-gfp* and *narJ(GGC)-gfp* strains (**Figure 3**, left panel).

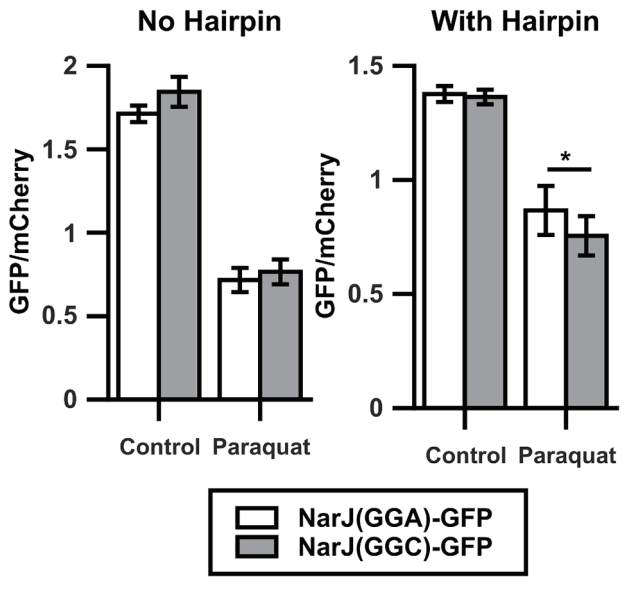

FIGURE 3 | Oxidative stress alters translation of Gly codons in natural context. Left graph shows the GFP/mCherry fluorescence ratio for *narJ* [NarJ(GGA)-GFP, white] or a mutant of the gene where a contiguous GGA pair was changed for a pair of GGC [NarJ(GGC)-GFP, gray] cloned in fusion to GFP. Shown in the right graph are similar experiments performed after introducing a hairpin that prevents ribosome sliding between the sequences coding for GFP and mCherry. *$p < 0.05$, two-tailed *t*-test ($n = 3$).

The sensitivity of our reporter may be decreased, if ribosomes that have translated *gfp* slide and initiate *mCherry* translation (70S-scanning initiation; Yamamoto et al., 2016), as then *mCherry* translation would not be completely independent of *gfp* translation. To avoid the possible effects of 70S-scanning, we introduced a sequence which is predicted to form a stable hairpin between both genes and has been previously shown to prevent ribosome sliding (Yamamoto et al., 2016). Using this construct, we observe a ~20% higher GFP/mCherry fluorescence ratio under oxidative stress when *narJ* is coded using GGA codons than when using GGC codons (**Figure 3**, right panel). This indicates that although under oxidative stress, GGA translation is less inhibited than translation of other Gly codons, and its effect on the amount of protein produced will strongly depend on the context where the codon is located.

tRNA^{Gly} Modulates the Response to Oxidative Stress

Alterations of translation efficiency of Gly codons probably alter level and/or cotranslational folding of several proteins under oxidative stress. In order to confirm that alterations in the active tRNA^{Gly} pool may affect the response to oxidative stress, we studied the response to oxidative stress in strains overproducing each tRNA^{Gly} isoacceptor from an IPTG inducible plasmid (pKK223-3). In all these experiments, the empty plasmid was used as a control for non-specific effects of the plasmid, and a plasmid coding for a tRNA that did not show changes in our aminoacylation experiments (tRNA$^{Tyr}_{GUA}$) was used as a control for the non-specific effects of tRNA overproduction. Overproduction of any of the tested tRNAs (including tRNA$^{Tyr}_{GUA}$) increased the sensitivity of carbohydrate fermentation to paraquat as measured by changes in the pH of culture media. Nevertheless, the effect was much stronger for the strain producing tRNA$^{Gly}_{CCC}$. Similarly, all the other tested phenotypes were also dependent on the overproduction of individual isoacceptors. For instance, while the stronger effect of overproduction of tRNA$^{Gly}_{GCC}$ was in preventing the loss of bacterial motility (measured as changes in colony diameter in low agar LB plates), overproduction of tRNA$^{Gly}_{UCC}$ mostly reduced culture yield (measured in 96-well plate cultures). As mentioned above, overproduction of tRNA$^{Gly}_{CCC}$ mainly increased the sensitivity of carbohydrate fermentation to paraquat (**Figure 4** and **Supplementary Figure S6**). In contrast to these phenotypes, cells overproducing these tRNAs did not show any change in their respective MIC for paraquat (25 μM in M9 media supplemented with branched amino acids).

In addition to generating idiosyncratic phenotypes, overproduction of each tRNA^{Gly} induced distinct changes in the aminoacylation of the other tRNA^{Gly} isoacceptors (**Figure 4D**). For instance, overproduction of tRNA$^{Gly}_{GCC}$ led to a ~10% decrease of the basal levels of aminoacylation of all tRNA^{Gly} isoacceptors. In contrast, tRNA$^{Gly}_{UCC}$ overproduction only altered basal aminoacylation of the other tRNA^{Gly} isoacceptors, but not its own. Finally, overproduction of tRNA$^{Gly}_{CCC}$ increased its basal level of aminoacylation while inducing a decrease in aminoacylation of other tRNA^{Gly} isoacceptors. Under oxidative stress, the effects were similarly diverse, with tRNA$^{Gly}_{GCC}$ inducing a decrease in tRNA$^{Gly}_{CCC}$ aminoacylation but tRNA$^{Gly}_{UCC}$ and,

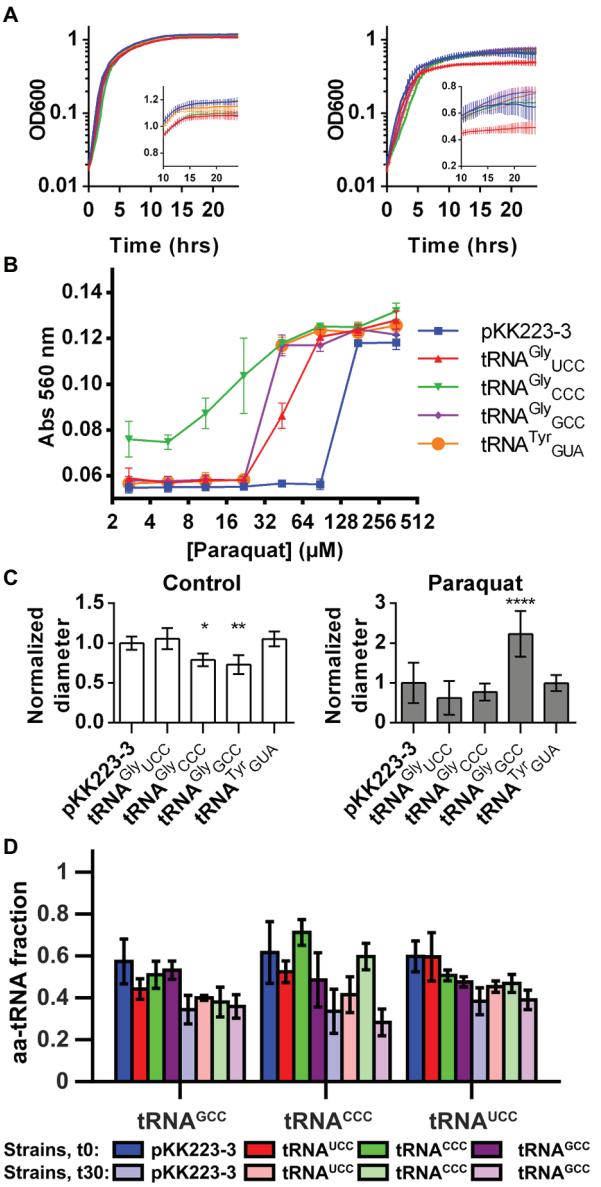

FIGURE 4 | Overproduction of tRNA^{Gly} alters the response of *E. coli* to oxidative stress. The effect of overproduction of diverse tRNA^{Gly} isoacceptors on the response to oxidative stress induced by paraquat was studied. **(A)** Effect of tRNA^{Gly} overproduction over growth curves performed at 37°C in control media (left) or media with paraquat (stress condition, right). Inset graph shows same data in linear scale for time points above 10 h. tRNA$^{Gly}_{CCC}$: green, tRNA$^{Gly}_{GCC}$: purple, tRNA$^{Gly}_{UCC}$: red, tRNA$^{Tyr}_{GUA}$: orange, and empty pKK223-3 plasmid: blue (n = 8). **(B)** Effect of tRNA^{Gly} overproduction on the fermentation of lactose as measured by changes in media pH after incubation at diverse paraquat concentrations. Higher absorbance indicates higher pH (lower fermentation of the carbohydrate). tRNA$^{Gly}_{CCC}$: green inverted triangles, tRNA$^{Gly}_{GCC}$: purple rhombus, tRNA$^{Gly}_{UCC}$: red triangles, tRNA$^{Tyr}_{GUA}$: orange circles, and empty pKK223-3 plasmid: blue squares (n = 3). **(C)** Effect of tRNA^{Gly} overproduction on bacterial motility as measured by changes in the diameter of colonies cultured on low agar LB plates in the absence (left) or presence (right) of paraquat. Data in graph represent the diameter of colonies normalized dividing by the average diameter of the control colonies. Each bar *(Continued)*

FIGURE 4 | represents the average of at least four replicates. $^{****}p \leq 0.0001$, $^{**}p \leq 0.01$, $^{*}p \leq 0.05$, one way ANOVA with Dunnett versus strain with empty plasmid at the corresponding condition. **(D)** Fraction of tRNA^{Gly} isoacceptors that is aminoacylated in strains overproducing diverse tRNA^{Gly} isoacceptors. Color represents the tRNA isoacceptor that is overproduced in strains where aminoacylation was quantified. tRNA$^{Gly}_{CCC}$: green, tRNA$^{Gly}_{GCC}$: purple, tRNA$^{Gly}_{UCC}$: red, empty pKK223-3 plasmid: blue. Darker colors correspond to measurements performed before paraquat addition, while light colors represent aminoacylation 30 min after addition of the stressor. Each bar represents the average of at least three replicates.

in particular, tRNA$^{Gly}_{CCC}$ inducing an increase in tRNA$^{Gly}_{CCC}$ aminoacylation level. Overall, under oxidative stress, the aminoacylation of the most abundant tRNA^{Gly} isoacceptor (tRNA$^{Gly}_{GCC}$) was the least sensitive to the levels of other tRNA^{Gly} isoacceptors, while tRNA$^{Gly}_{CCC}$ showed the highest sensitivity.

In summary, the specificity of the phenotypes induced by overproduction of each tRNA^{Gly} isoacceptor and the fact that most were observed under oxidative stress but not during normal growth suggests that these effects are triggered by individual tRNA^{Gly} isoacceptors and are not simply secondary effects of tRNA overproduction. Nevertheless, it is currently not possible to determine a cause-effect relation between particular tRNAs and phenotypes because changes in the levels of any tRNA^{Gly} isoacceptor alter the aminoacylation of other tRNAs.

DISCUSSION

Regulation of Translation by Changes in tRNA Concentration

Here, we demonstrate that oxidative stress induces specific alterations in the tRNA^{Gly} pools and concurrent changes in Gly codon translation rates. Nevertheless, changes in protein levels are only observed under strong oxidative stress, low Gly availability, and particular genetic context. Taken together, our findings suggest that diversity of codon translation speeds is only observed when the speed of Gly codon translation is similar to or slower than translation initiation which is usually considered the limiting step of the complete process (Hersch et al., 2014; Subramaniam et al., 2014). Thus, it appears that the translational response to oxidative stress is complex and effects on elongation are only observed under the most hostile conditions. Either lower paraquat concentration or higher Gly availability induces a condition where changes in *gfp* codons do not affect protein production.

In agreement with our observations, reports of experiments performed in other *E. coli* strains also indicate that the relevance of elongation in determining the overall speed of protein production increases at stronger oxidative stress conditions, in this case induced by higher concentrations of H_2O_2 (Zhu and Dai, 2019). In contrast to what we have observed, but in agreement with previous reports (Zhong et al., 2015), the authors of these experiments observe a decrease in the concentration of all tRNAs. Thus, the translational response to oxidative stress seems to be strain dependent. It is currently

difficult to predict what determines these different behaviors between *E. coli* strains. For example, the lack of a toxin-antitoxin system or a smaller nuclease activity in the strain we used could prevent tRNA cleavage under oxidative stress. Alternatively, a decreased protease activity or increased amino acid usage might produce stronger limitations of amino acid availability in strains used by other groups, thereby inducing degradation of tRNAs (Svenningsen et al., 2017).

Further research will be required to clarify the peculiarities that induce different behaviors between *E. coli* strains. Nevertheless, our data show that in *E. coli* K-12 MG1655, an important model strain, only tRNA^{Gly} is inactivated under oxidative stress. Although all isoacceptors are deacylated, the decrease in the speed of GGA translation is smaller than observed for other Gly codons. This should allow preferential translation of GGA enriched genes under oxidative stress. As mentioned previously, GGA is used with a lower frequency than other Gly codons in *E. coli* K-12 MG1655 (**Table 1**; **Supplementary Figure S2**; **Supplementary Table S1**). Only 79 genes (~1.8% of all *E. coli* K-12 MG1655 genes) use GGA as 3% or more of their codons. Many of the genes that use GGA codons with high frequency are implicated in the negative regulation of cell growth and cellular defense, including toxins from three toxin-antitoxin systems (*chpB*, *mazF*, and *ralR*) that might explain the differences observed in growth after overproduction of tRNA$^{Gly}_{UCC}$. The high abundance of GGA codons in genes such as *rmf*, which is involved in ribosome hibernation during stationary phase or *yciH* that binds the ribosome and alters the expression of stress response genes, and growth in minimal medium could further explain some of the observed growth phenotypes. Also, the enrichment of genes implicated in cell adhesion or motility (*yraK*, *ydeQ*, *yadK*, *flhE*, and *chaC*) could explain different motility behaviors between strains overexpressing the different tRNA^{Gly} genes. A comprehensive list of functions enriched in the list of genes with high GGA codon usage is given in **Supplementary Table S2**. Nevertheless, care should be taken when extrapolating results obtained by using fluorescent reporters to these GGA-rich genes, as our results indicate that sensitivity to oxidative stress may strongly depend on the context where GGA codons are located. This context sensitivity may derive from different speeds of translation depending on neighboring codons (Chevance et al., 2014), for instance, due to interactions with other tRNAs in the ribosome (Buchan et al., 2006). As we are analyzing translation of full-length genes and not isolated GGA codons, different sensitivities may additionally arise from a slow translation initiation or specific patterns of amino acids or codons where translation elongation is very slow, making any change in GGA translation undetectable due to other limiting steps in translation (Hersch et al., 2014). As mentioned previously, this might explain the lack of differences between translation of each Gly codon under some conditions. Oxidation of ribosomal RNA (rRNA) or proteins (Katz and Orellana, 2012; Liu et al., 2012; Willi et al., 2018) might explain such a decrease in translation initiation. In this context, the fact that the anti-Shine-Dalgarno of *E. coli* presents repeats (ACCUCC) of the tRNA^{Gly} anticodon sequences (NCC) might suggest a similar modification mechanism. Nevertheless, the little available data (Willi et al., 2018) suggests that the 3' extreme is more resistant to oxidative stress derived modifications than other segments of the rRNA.

Based on current data, it is difficult to determine why translation of GGA is less inhibited than translation of the other Gly codons. One possibility is that there is a lower requirement of aa-tRNA$^{Gly}_{UCC}$ as a consequence of the low frequency of usage of GGA codons (**Table 1**; **Supplementary Figure S2**). If true, this would limit the sensitivity of translation to changes in the levels of active tRNA$^{Gly}_{UCC}$. Nevertheless, relations between tRNA and translation appear to be complex, and the effect of changes of a single tRNA might be difficult to predict. For instance, while we observe that increasing the concentration of tRNA$^{Gly}_{GCC}$ may result in significant alterations in the levels of aminoacylation of the other tRNA^{Gly} isoacceptors (**Figure 4D**), we also found that at least under some conditions, a decrease in concentration of the same tRNA can have very limited effects on translation of the codons not directly translated by the affected tRNA (**Supplementary Figure S5**).

Final Remarks

Recent research on the effects of oxidative stress on the bacterial translation machinery has shown a very diverse set of effects ranging from changes in error rates (Ling and Söll, 2010; Wu et al., 2014; Steiner et al., 2019) to translation inhibition (Kojima et al., 2009; Nagano et al., 2015; Zhong et al., 2015; Yutthanasirikul et al., 2016; Zhu and Dai, 2019). Such results appear to derive from alterations to all components of the machinery such as changes in tRNA (Katz and Orellana, 2012; Liu et al., 2012; Zhong et al., 2015; Zhu and Dai, 2019), aminoacyl-tRNA synthetases (Ling and Söll, 2010; Katz and Orellana, 2012; Wu et al., 2014; Steiner et al., 2019), ribosomes (Katz and Orellana, 2012; Liu et al., 2012; Willi et al., 2018), and elongation factors (Kojima et al., 2009; Katz and Orellana, 2012; Nagano et al., 2015; Yutthanasirikul et al., 2016). Nevertheless, comparison of our results with these previous reports indicates that the relevance of each of these changes to bacterial adaptation depends not only on the culture conditions but also on the strains being analyzed. Thus, further studies will be required to understand the relevance of alterations in each component of the translation machinery in the adaptation to diverse degrees of stress and how the genetic background of each strain determines this response.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

Aminoacylation experiments were performed by AK and LL. tRNA level quantifications were performed by AP and LL. Phenotypic characterization was performed by LL and AP. tRNA purification was performed by VB, mass spectrometry

experiments by SK, and mass spectrometry data analyses by SK and JL. Cloning, mutagenesis, and GFP/mCherry analyses were performed by SE, DG, and LL. Codon enrichment analysis was performed by AK. MI, AK, and SL supervised experiments and interpreted the data. AK wrote the paper with contributions from all authors. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by Vicerrectoría de Investigación y Desarrollo (VID), Universidad de Chile (ENL05/18 to AK); Fondo Nacional de Desarrollo Científico y Tecnológico (1191074 to AK); Comisión Nacional de Investigación Científica y Tecnológica (79130044 to AK); National Institutes of Health Grant (GM65183 to MI); ERC Starting Grant (ERC-2012-StG 310489-tRNAmodi to SL); Swiss National Science Foundation (NCCR RNA & Disease to SL); and a CONICYT Doctorado Nacional scholarship from Comisión Nacional de Investigación Científica y Tecnológica (21151441 to LL).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00856/full#supplementary-material

REFERENCES

Alings, F., Sarin, L. P., Fufezan, C., Drexler, H. C. A., and Leidel, S. A. (2015). An evolutionary approach uncovers a diverse response of TRNA 2-thiolation to elevated temperatures in yeast. *RNA* 21, 202–212. doi: 10.1261/rna.048199.114

Bald, T., Barth, J., Niehues, A., Specht, M., Hippler, M., and Fufezan, C. (2012). pymzML--Python module for high-throughput bioinformatics on mass spectrometry data. *Bioinformatics* 28, 1052–1053. doi: 10.1093/bioinformatics/bts066

Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., et al. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science* 277, 1453–1462. doi: 10.1126/science.277.5331.1453

Blondel, C. J., Jiménez, J. C., Leiva, L. E., Alvarez, S. A., Pinto, B. I., Contreras, F., et al. (2013). The type VI secretion system encoded in *Salmonella* pathogenicity island 19 is required for *Salmonella enterica* serotype gallinarum survival within infected macrophages. *Infect. Immun.* 81, 1207–1220. doi: 10.1128/IAI.01165-12

Boccaletto, P., Machnicka, M. A., Purta, E., Piatkowski, P., Baginski, B., Wirecki, T. K., et al. (2017). MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res.* 46, D303–D307. doi: 10.1093/nar/gkx1030

Böck, A., and Neidhardt, F. C. (1966). Location of the structural gene for Glycyl ribonucleic acid synthetase by means of a strain of *Escherichia coli* possessing an unusual enzyme. *Z. Vererbungsl.* 98, 187–192. doi: 10.1007/BF00888946

Buchan, J. R., Aucott, L. S., and Stansfield, I. (2006). tRNA properties help shape codon pair preferences in open reading frames. *Nucleic Acids Res.* 34, 1015–1027. doi: 10.1093/nar/gkj488

Chambers, M. C., Maclean, B., Burke, B., Amodei, D., Ruderman, D. L., Neumann, S., et al. (2012). A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* 30, 918–920. doi: 10.1038/nbt.2377

Chan, P. P., and Lowe, T. M. (2009). GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Res.* 37(Suppl. 1), D93–D97. doi: 10.1093/nar/gkn787

Charbon, G., Bjørn, L., Mendoza-Chamizo, B., Frimodt-Møller, J., and Løbner-Olesen, A. (2014). Oxidative DNA damage is instrumental in hyperreplication stress-induced inviability of *Escherichia coli*. *Nucleic Acids Res.* 42, 13228–13241. doi: 10.1093/nar/gku1149

Chevance, F. F. V., Le Guyon, S., and Hughes, K. T. (2014). The effects of codon context on in vivo translation speed. *PLoS Genet.* 10:e1004392. doi: 10.1371/journal.pgen.1004392

Chiang, S. M., and Schellhorn, H. E. (2012). Regulators of oxidative stress response genes in *Escherichia coli* and their functional conservation in bacteria. *Arch. Biochem. Biophys.* 525, 161–169. doi: 10.1016/j.abb.2012.02.007

Cohen, N. R., Lobritz, M. A., and Collins, J. L. (2013). Microbial persistence and the road to drug resistance. *Cell Host Microbe* 13, 632–642. doi: 10.1016/j.chom.2013.05.009

Datsenko, K. A., and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* 97, 6640–6645. doi: 10.1073/pnas.120163297

Dong, H., Nilsson, L., and Kurland, C. G. (1996). Co-variation of tRNA abundance and codon usage in *Escherichia coli* at different growth rates. *J. Mol. Biol.* 260, 649–663. doi: 10.1006/jmbi.1996.0428

Elgamal, S., Katz, A., Hersch, S. J., Newsom, D., White, P., Navarre, W. W., et al. (2014). EF-P dependent pauses integrate proximal and distal signals during translation. *PLoS Genet.* 10:e1004553. doi: 10.1371/journal.pgen.1004553

Folk, W. R., and Berg, P. (1970a). Characterization of altered forms of Glycyl transfer ribonucleic acid synthetase and the effects of such alterations on aminoacyl transfer ribonucleic acid synthesis in vivo. *J. Bacteriol.* 102, 204–212.

Folk, W. R., and Berg, P. (1970b). Isolation and partial characterization of *Escherichia coli* mutants with altered Glycyl transfer ribonucleic acid synthetases. *J. Bacteriol.* 102, 193–203.

Gambino, M., and Cappitelli, F. (2016). Mini-review: biofilm responses to oxidative stress. *Biofouling* 32, 167–178. doi: 10.1080/08927014.2015.1134515

Ha, D., Kuchma, S. L., and O'Toole, G. A. (2014). Plate-based assay for swimming motility in *Pseudomonas aeruginosa*. *Methods Mol. Biol.* 1149, 59–65. doi: 10.1007/978-1-4939-0473-0_7

Hassan, H. M., and Fridovich, I. (1979). Paraquat and *Escherichia coli*. Mechanism of production of extracellular superoxide radical. *J. Biol. Chem.* 254, 10846–10852.

Henriques, M. X., Catalão, M. J., Figueiredo, J., Gomes, J. P., and Filipe, S. R. (2013). Construction of improved tools for protein localization studies in *Streptococcus pneumoniae*. *PLoS One* 8:e55049. doi: 10.1371/journal.pone.0055049

Hersch, S. J., Elgamal, S., Katz, A., Ibba, M., and Navarre, W. W. (2014). Translation initiation rate determines the impact of ribosome stalling on bacterial protein synthesis. *J. Biol. Chem.* 289, 28160–28171. doi: 10.1074/jbc.M114.593277

Imlay, J. A. (2008). Cellular defenses against superoxide and hydrogen peroxide. *Annu. Rev. Biochem.* 77, 755–776. doi: 10.1146/annurev.biochem.77.061606.161055

Imlay, J. A. (2013). The molecular mechanisms and physiological consequences of oxidative stress: lessons from a model bacterium. *Nat. Rev. Microbiol.* 11, 443–454. doi: 10.1038/nrmicro3032

Katz, A., Elgamal, S., Rajkovic, A., and Ibba, M. (2016). Non-canonical roles of tRNAs and tRNA mimics in bacterial cell biology. *Mol. Microbiol.* 101, 545–558. doi: 10.1111/mmi.13419

Katz, A., and Orellana, O. (2012). "Protein synthesis and the stress response" in *Cell-free protein synthesis*. ed. M. Biyani (Rijeka, Croatia: IntechOpen), 111–134.

Keseler, I. M., Mackie, A., Santos-Zavaleta, A., Billington, R., Bonavides-Martínez, C., Caspi, R., et al. (2017). The EcoCyc database: reflecting new knowledge about *Escherichia coli* K-12. *Nucleic Acids Res.* 45, D543–D550. doi: 10.1093/nar/gkw1003

Kojima, K., Motohashi, K., Morota, T., Oshita, M., Hisabori, T., Hayashi, H., et al. (2009). Regulation of translation by the redox state of elongation factor G in the cyanobacterium *Synechocystis* sp. PCC 6803. *J. Biol. Chem.* 284, 18685–18691. doi: 10.1074/jbc.M109.015131

Leufken, J., Niehues, A., Sarin, L. P., Wessel, F., Hippler, M., Leidel, S. A., et al. (2017). pyQms enables universal and accurate quantification of mass spectrometry data. *Mol. Cell. Proteomics* 16, 1736–1745. doi: 10.1074/mcp.M117.068007

Lewis, K. (2010). Persister cells. *Annu. Rev. Microbiol.* 64, 357–372. doi: 10.1146/annurev.micro.112408.134306

Ling, J., and Söll, D. (2010). Severe oxidative stress induces protein mistranslation through impairment of an aminoacyl-tRNA synthetase editing site. *Proc. Natl. Acad. Sci. U. S. A.* 107, 4028–4033. doi: 10.1073/pnas.1000315107

Liu, M., Gong, X., Alluri, R. K., Wu, J., Sablo, T., and Li, Z. (2012). Characterization of RNA damage under oxidative stress in *Escherichia coli*. *Biol. Chem.* 393, 123–132. doi: 10.1515/hsz-2011-0247

Martens, L., Chambers, M., Sturm, M., Kessner, D., Levander, F., Shofstahl, J., et al. (2011). mzML--a community standard for mass spectrometry data. *Mol. Cell. Proteomics* 10:R110.000133. doi: 10.1074/mcp.R110.000133

Nagano, T., Yutthanasirikul, R., Hihara, Y., Hisabori, T., Kanamori, T., Takeuchi, N., et al. (2015). Oxidation of translation factor EF-G transiently retards the translational elongation cycle in *Escherichia coli*. *J. Biochem.* 158, 165–172. doi: 10.1093/jb/mvv026

Nguyen, G. T., Green, E. R., and Mecsas, J. (2017). Neutrophils to the ROScue: mechanisms of NADPH oxidase activation and bacterial resistance. *Front. Cell. Infect. Microbiol.* 7:373. doi: 10.3389/fcimb.2017.00373

Raczniak, G., Becker, H. D., Min, B., and Söll, D. (2001). A single amidotransferase forms asparaginyl-tRNA and glutaminyl-tRNA in *Chlamydia trachomatis*. *J. Biol. Chem.* 276, 45862–45867. doi: 10.1074/jbc.M109494200

Rojas, J., Castillo, G., Leiva, L. E., Elgamal, S., Orellana, O., Ibba, M., et al. (2018). Codon usage revisited: lack of correlation between codon usage and the number of tRNA genes in enterobacteria. *Biochem. Biophys. Res. Commun.* 502, 450–455. doi: 10.1016/j.bbrc.2018.05.168

Rui, B., Shen, T., Zhou, H., Liu, J., Chen, J., Pan, X., et al. (2010). A systematic investigation of *Escherichia coli* central carbon metabolism in response to superoxide stress. *BMC Syst. Biol.* 4:122. doi: 10.1186/1752-0509-4-122

Salazar, J. C., Ambrogelly, A., Crain, P., McCloskey, J. A., and Söll, D. (2004). A truncated aminoacyl-tRNA synthetase modifies RNA. *Proc. Natl. Acad. Sci. U. S. A.* 101, 7536–7541. doi: 10.1073/pnas.0401982101

Sarin, L. P., Kienast, S. D., Leufken, J., Ross, R., Dziergowska, A., Debiec, K., et al. (2018). Nano LC-MS using capillary columns enables accurate quantification of modified ribonucleosides at low femtomol levels. *RNA* 24, 1403–1417. doi: 10.1261/rna.065482.117

Shen, T., Rui, B., Zhou, H., Zhang, X., Yi, Y., Wen, H., et al. (2013). Metabolic flux ratio analysis and multi-objective optimization revealed a globally conserved and coordinated metabolic response of *E. coli* to paraquat-induced oxidative stress. *Mol. BioSyst.* 9, 121–132. doi: 10.1039/C2MB25285F

Shimizu, K. (2016). Metabolic regulation and coordination of the metabolism in bacteria in response to a variety of growth conditions. *Adv. Biochem. Eng. Biotechnol.* 155, 1–54. doi: 10.1007/10_2015_320

Slauch, J. M. (2011). How does the oxidative burst of macrophages kill bacteria? Still an open question. *Mol. Microbiol.* 80, 580–583. doi: 10.1111/j.1365-2958.2011.07612.x

Starzyk, R. (1984). tRNA base modifications and gene regulation. *Trends Biochem. Sci.* 9, 333–334. doi: 10.1016/0968-0004(84)90053-7

Steiner, R. E., Kyle, A. M., and Ibba, M. (2019). Oxidation of phenylalanyl-tRNA synthetase positively regulates translational quality control. *Proc. Natl. Acad. Sci. U. S. A.* 116, 10058–10063. doi: 10.1073/pnas.1901634116

Subramaniam, A. R., Deloughery, A., Bradshaw, N., Chen, Y., O'Shea, E., Losick, R., et al. (2013a). A serine sensor for multicellularity in a bacterium. *eLife* 2:e01501. doi: 10.7554/eLife.01501

Subramaniam, A. R., Pan, T., and Cluzel, P. (2013b). Environmental perturbations lift the degeneracy of the genetic code to regulate protein levels in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 110, 2419–2424. doi: 10.1073/pnas.1211077110

Subramaniam, A. R., Zid, B. M., and O'Shea, E. K. (2014). An integrated approach reveals regulatory controls on bacterial translation elongation. *Cell* 159, 1200–1211. doi: 10.1016/j.cell.2014.10.043

Svenningsen, S. L., Kongstad, M., Stenum, T. S., Muñoz-Gómez, A. J., and Sørensen, M. A. (2017). Transfer RNA is highly unstable during early amino acid starvation in *Escherichia coli*. *Nucleic Acids Res.* 45, 793–804. doi: 10.1093/nar/gkw1169

Wiegand, I., Hilpert, K., and Hancock, R. E. W. (2008). Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nat. Protoc.* 3, 163–175. doi: 10.1038/nprot.2007.521

Willi, J., Küpfer, P., Evéquoz, D., Fernandez, G., Katz, A., Leumann, C., et al. (2018). Oxidative stress damages rRNA inside the ribosome and differentially affects the catalytic center. *Nucleic Acids Res.* 46, 1945–1957. doi: 10.1093/nar/gkx1308

Wu, J., Fan, Y., and Ling, J. (2014). Mechanism of oxidant-induced mistranslation by threonyl-tRNA synthetase. *Nucleic Acids Res.* 42, 6523–6531. doi: 10.1093/nar/gku271

Wu, Y., Vulić, M., Keren, I., and Lewis, K. (2012). Role of oxidative stress in persister tolerance. *Antimicrob. Agents Chemother.* 56, 4922–4926. doi: 10.1128/AAC.00921-12

Yamamoto, H., Wittek, D., Gupta, R., Qin, B., Ueda, T., Krause, R., et al. (2016). 70S-scanning initiation is a novel and frequent initiation mode of ribosomal translation in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 113, E1180–E1189. doi: 10.1073/pnas.1524554113

Yutthanasirikul, R., Nagano, T., Jimbo, H., Hihara, Y., Kanamori, T., Ueda, T., et al. (2016). Oxidation of a cysteine residue in elongation factor EF-Tu reversibly inhibits translation in the cyanobacterium *Synechocystis* sp. PCC 6803. *J. Biol. Chem.* 291, 5860–5870. doi: 10.1074/jbc.M115.706424

Zhong, J., Xiao, C., Gu, W., Du, G., Sun, X., He, Q., et al. (2015). Transfer RNAs mediate the rapid adaptation of *Escherichia coli* to oxidative stress. *PLoS Genet.* 11:e1005302. doi: 10.1371/journal.pgen.1005302

Zhu, M., and Dai, X. (2019). Maintenance of translational elongation rate underlies the survival of *Escherichia coli* during oxidative stress. *Nucleic Acids Res.* 47, 7592–7604. doi: 10.1093/nar/gkz467

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: info@frontiersin.org | +41 21 510 17 00

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership